

86257

STIC-Biotech/ChemLib

Fr m: Bui, Phuong
Sent: Friday, February 07, 2003 12:00 PM
To: STIC-Biotech/ChemLib
Subj ct: 09/857612 sequence search request

Please search in both commercial and interference databases the following sequences:

- 1) n.a. SEQ ID NO:13
- 2) a.a. SEQ ID NO:14
- 3) DNA encoding SEQ ID NO:14

Thank You. Phuong.

Phuong Bui
305-1996
CM1 Office 9A09, Mailbox 9E12
AU 1638

Point of Contact:
Barb O'Bryen
Technical Information Specialist
STIC CM1 6A05 308-4291

Searcher: Bob
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 2-20-03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

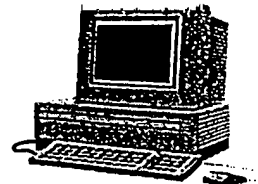
TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

BioTech-Chem Library

Search Results

Feedback Form (Optional)



Scientific & Technical Information Center

The search results generated for your recent request are attached. If you have any questions or comments (compliments or complaints) about the scope or the results of the search, please contact *the BioTech-Chem searcher* who conducted the search *or contact*:

Mary Hale, Supervisor, 308-4258
CM-1 Room 1E01

Voluntary Results Feedback Form

➤ *I am an examiner in Workgroup:* (Example: 1610)

➤ *Relevant prior art found, search results used as follows:*

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ *Relevant prior art not found:*

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Search results were not useful in determining patentability or understanding the invention.

Other Comments:

Drop off completed forms at the Circulation Desk CM-1, or send to Mary Hale, CM1-1E01 or mary.hale@uspto.gov

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 11:10:36 ; Search time 2449 seconds
(without alignments)
15828.884 Million cell updates/sec

Title: US-09-857-612a-13

Perfect score: 1332

Sequence: 1 atgaagaagaacacagaaga.....gtaattgttttcgsggttaa 1332

Scoring table: IDENTITY_NTC

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBankl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pin:*

35: em_htg_rod:*

36: em_htg_mem:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	449.6	33.8	1299	6 AX090309	AX090309 Sequence
2	449.6	33.8	1299	8 AX133614	AX133614 Arabidops
3	449.6	33.8	1457	8 AF367326	AF367326 Arabidops
4	448	33.6	1498	8 AY087433	AY087433 Arabidops
5	369.4	27.7	833	8 AF209909	AF209909 Prunus du
6	298.2	22.4	1562	6 AX037586	AX037586 Sequence
7	298.2	22.4	1562	6 AX037605	AX037605 Sequence
8	298.2	22.4	99690	8 AC004557	AC004557 Genomic s
9	249	18.7	264	6 AX090319	AX090319 Sequence
10	206	15.5	273	6 AX090320	AX090320 Sequence
11	197.6	14.8	134188	2 AC120983	AC120983 Oryza sat
12	125	9.4	123620	8 AC103891	AC103891 Oryza sat
13	122.8	9.2	139534	2 AP005640	AP005640 Oryza sat
14	121.6	9.1	147348	8 AP004073	AP004073 Oryza sat
15	96.2	7.2	128210	2 AP005643	AP005643 Oryza sat
16	91.4	6.9	246	11 AL773076	AL773076 Arabidops
17	88.2	6.6	1346	10 RN062803	u62803 Rattus norv
18	87	6.5	147203	8 AP003687	AP003687 Oryza sat
19	86.6	6.5	1362	10 RN062803	X54096 Rat mRNA fo
20	79.6	6.0	1341	10 MUSLCATX	J05154 Murine Chol
21	79.6	6.0	1354	10 BC028861	BC028861 Mus muscu
22	76.8	5.8	1155	6 AX090327	AX090327 Sequence
23	76.6	5.8	1338	4 AF272861	AF272861 Tupai a g1
24	65.8	4.9	1341	4 RABLCAT	D13668 Oryctolagus
25	65.8	4.9	134374	2 AC121465	AC121465 Rattus no
26	65.8	4.9	251055	2 AC094385	AC094385 Rattus no
27	64.6	4.8	1297	9 HUMLCATG	M26268 Human lecit
28	64.6	4.8	1355	6 HUMLCATG	M17959 Human lecit
29	64.6	4.8	1307	9 HUMLCATG	E01185 DNA encodin
30	64.6	4.8	1360	9 BC014781	BC014781 Homo sapi
31	64.6	4.8	1447	9 HSLCAT1	X06537 Human mRNA
32	64.6	4.8	1744	6 AX077318	AX077318 Sequence
33	64.6	4.8	1744	6 AX409598	AX409598 Sequence
34	64.6	4.8	1744	9 HUMLCAT	M12625 Human lecit
35	64.6	4.8	4308	6 AX077320	AX077320 Sequence
36	64	4.8	1137	6 E26768	E26768 Novel prote
37	64	4.8	1236	6 E26773	E26773 Novel prote
38	64	4.8	1239	10 AF468958	AF468958 Mus muscu
39	64	4.8	2721	10 BC019373	BC019373 Mus muscu
40	61.2	4.6	73944	3 AC006402	AC006402 Drosophi1
41	61.2	4.6	182901	3 AC093198	AC093198 Drosophi1
42	61.2	4.6	192132	3 AC093440	AC093440 Drosophi1
43	61.2	4.6	309928	3 AE003666	AE003666 Drosophi1
44	60.4	4.5	1137	6 E26766	E26766 Novel prote
45	60.4	4.5	1146	6 E26769	E26769 Novel prote

ALIGNMENTS

RESULT 1

AX090309 1299 bp DNA linear PAT 21-MAR-2001

LOCUS AX090309

DEFINITION Sequence 2 from Patent WO0116308.

ACCESSION AX090309

VERSION AX090309.1 GI:13444179

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana

thale cress.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

AUTHORS Lasserre M. and van Eenennaam A.

TITLE Plant sterol acyltransferases

Pred. No. is the number of results predicted by chance to have a

DEFINITION Arabidopsis thaliana clone 35408 mRNA, complete sequence.
ACCESSION AY087433
VERSION AY087433.1 GI:21406157
KEYWORDS FLI.CDNA
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS 1 (bases 1 to 1498)
Haas,B.J., Volkovskiy,N., Town,C.D., Troupkan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
TITLE Full-length messenger RNA sequences greatly improve genome annotation
JOURNAL Genome Biol. (2002) In press
REFERENCE 2 (bases 1 to 1498)
AUTHORS Brover,V., Troupkan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
TITLE Full-length cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1498)
AUTHORS Brover,V., Troupkan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA
COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants; including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the WS or Ler ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.
FEATURES
source location/Qualifiers.
1..1498
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="35408"
82..1380
/codon_start=-1
/product="unknown"
/protein_id="AA064980.1"
/db_xref="GI:21593031"
/translation="MKRISHYEVVIAILVVTMTSMCAVGSNVPYLILVPGNSNQ LEVRLDREYKSPSSVMSWLYPIHKSGWFLMFDAVILSPTRCFSRMMLYDP DLDDYONAPGVOTRPHFEGSTSLILYDLPLRDAITSNEHLVALLEKCGVNDOTIL GAYDERYVGAASGHSRVSASOPLDLODKOIVERTSSNEKHPVLLISHSIGLPVLEF LNRTPSSMRKRYIKHYALAAPMGCTISOKTKRASGTLVLPVLPNPLVLRHQTSES NQWLPLSTKVFHDKRLELVVTPQVNTAYVMDRFADIGSQGVAPVKTVLPITTEL MTPGVPTCYIGRGVDPPEVLAMGKGFDPKPKYDGGTVMASIALAKVDSLNT VEIDGVSHTEILDEIALKIMKOISIINELANVNAVNE"
BASE COUNT 425 a 322 c 343 g 408 t
ORIGIN
Query Match 33.6%; Score 448; DB 8; Length 1498;
Best Local Similarity 62.3%; Pred. No. 2.3e-108;
Matches 785; Conservative 0; Mismatches 445; Indels 30; Gaps 4;

Dh 171 TGTGTACCCCTTTGATTGTGTGCTTCCAGAAACGAGGATACCAAGCTAGAGGTACGGCTGGA 230
Qy 165 CAATCAGTACAAAGCCCTTACTTTATCTGCGAATCATAGTACCCCTCATCAAGAAAA 224
Dh 221 CAGAGAAATACAAAGCCAGTAGTGTGTGAGCAGCTGTATATCCGATTCATTAAGAA 290
Qy 225 GA---ATGAGATGTCAGACTTGTGTGATTCAGAGTGCATACACTTGGCTTCACTCA 281
Dh 291 GAGTGTGATAGTGTATAGCTATGTTGATGATCAGACAGTGTATATGTCTCCCTTAC 350
Qy 282 ATGCTTTCGCCGAGCATGACCTTCATTAACCCCAAGAACTCGATGATTACTTCAACAC 341
Dh 351 GTCCTTACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 410
Qy 342 TCCCTGGGTTGAGACCGGGTCCCTCATTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 401
Dh 411 TCCCTGGTTCCAAAACCGGGTCCCTCATTGTTGTTGTTGTTGTTGTTGTTGTTGTT 470
Qy 402 TCCCTGGTTCCAAAACCGGGTCCCTCATTGTTGTTGTTGTTGTTGTTGTTGTTGTT 458
Dh 471 CCTCTGCTCCGAGATGCCATCTTACATGATGATGATGATGATGATGATGATGATG 530
Qy 459 GCTTGGCTACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 518
Dh 531 ATGCGGTATGTTATACGAACCAACATCCCTAGACCTCATATGATTTGAGTACGGCT 590
Qy 519 AGCTCTAGAGTACACCTTTCACAGTGGGTTCCAAAGTTCTTCAAGATCTAAGCAATTT 578
Dh 591 GCTCTCTTCGGCCACCGGCTCCGCTGATGACCTCAGAGTCTTCCAGAACCTTCAAA 650
Qy 579 GATTAAGAGAGCAACCAATTCATTAATGAGAGCAGATGATGATGATGATGATGATG 638
Dh 651 GGTGAAAGAACTTACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 710
Qy 639 AGAGAGCTATTTGTCTTCACTACACTTAATAGAAACCCCTTGGCGCAAAATTT 698
Dh 711 AGAGAGCTTTTGTCTTCACTACACTTAATAGAAACCCCTTGGCGCAAAATTT 770
Qy 699 CATCAAACTTATGCTCTTTCAGCTCAGTGGGTTGTTGTTGTTGTTGTTGTTGTTG 758
Dh 771 CATCAAACTTATGCTCTTTCAGCTCAGTGGGTTGTTGTTGTTGTTGTTGTTGTTG 830
Qy 759 CTTTCATCTGGCAGACCTTTGGAGTGGGCTTATGAGGAGGATGATGAGGATGATG 818
Dh 831 ATTTCCTTCTGGCAGACCTTTGGAGTGGGCTTATGAGGAGGATGATGAGGATGATG 890
Qy 819 ACAAGAGAGCTCGAGAGTAACTTGGCTTTGGCTTAAACCCCAAAATTTT---GATC 875
Dh 891 TCAGAGAGAGCTCGAGAGTAACTTGGCTTTGGCTTAAACCCCAAAATTTT---GATC 950
Qy 876 TCAGAGAGAGCTCGAGAGTAACTTGGCTTTGGCTTAAACCCCAAAATTTT---GATC 935
Dh 951 AACTTAACCGCTTGTGTAAGTCCCAAGGATTTAACTACACAGCTTACGAGATGATG 1010
Qy 936 TCTAAGAGAGCTTGTGTAAGTCCCAAGGATTTAACTAAGAGAGGATTTAACTAAG 995
Dh 1011 TTTTCAAGAGAGCTTGTGTAAGTCCCAAGGATTTAACTAAGAGAGGATTTAACTA 1070
Qy 996 GATAGAGAGAGCTTGTGTAAGTCCCAAGGATTTAACTAAGAGAGGATTTAACTAAG 1055
Dh 1071 AAGAGAGAGAGCTTGTGTAAGTCCCAAGGATTTAACTAAGAGAGGATTTAACTAAG 1130
Qy 1056 AACCTTGAAGAGCTTGTGTAAGTCCCAAGGATTTAACTAAGAGAGGATTTAACTAAG 1115
Dh 1131 TACACCGAGAGCTTGTGTAAGTCCCAAGGATTTAACTAAGAGAGGATTTAACTAAG 1190
Qy 1116 TGGGAGAGAGCTTGTGTAAGTCCCAAGGATTTAACTAAGAGAGGATTTAACTAAG 1175
Dh 1191 TGGAGAGAGAGCTTGTGTAAGTCCCAAGGATTTAACTAAGAGAGGATTTAACTAAG 1235
Qy 1176 AGAGAGAGAGAGCTTGTGTAAGTCCCAAGGATTTAACTAAGAGAGGATTTAACTAAG 1235
Dh 1236 -----GAAAGTCAATACCTTGAACACCGTAGAGATGATGAGATGATGATGATGAT 1289

Result 5	AF209909	833 bp	DNA	Linear	PLN_03-JAN-2000
LOCUS	AF209909				
DEFINITION	Prunus dulcis lecithin-cholesterol acyl transferase (LCAT1) gene,				
ACCESSION	AF209909				
VERSION	AF209909.1				
KEYWORDS	GI:6653737				
SOURCE	Prunus dulcis.				
ORGANISM	Prunus dulcis				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.				
AUTHORS	1 (bases 1 to 833)				
TITLE	Ma,R.C. and Oliveira,M.M.				
JOURNAL	Molecular Cloning of A Lecithin-Cholesterol Acyl Transferase Homolog Gene from Almond (Prunus dulcis)				
REFERENCE	Unpublished				
AUTHORS	2 (bases 1 to 833)				
TITLE	Ma,R.C. and Oliveira,M.M.				
JOURNAL	Direct Submission				
FEATURES	Submitted (01-DEC-1999) Lab of Plant Genetic Engineering, Instituto de Biologia Experimental e Tecnologica, Quinta do Marquês, Oeiras/Lisboa, Lisboa Apct.12-2/80, Portugal				
SOURCE	Location/Qualifiers				
gene	1..833				
gene	/organism="Prunus dulcis"				
gene	/cultivar="Boa Casta"				
gene	/db_xref="taxon:3755"				
gene	<1..>802				
gene	/gene="LCAT1"				
gene	<1..>802				
gene	/gene="LCAT1"				
gene	/product="lecithin-cholesterol acyl transferase"				
gene	<1..802				
gene	/gene="LCAT1"				
gene	/product="lecithin-cholesterol acyl transferase"				
gene	<1..802				
gene	/gene="LCAT1"				
gene	/note="phosphatidylcholine-sterol acyltransferase"				
gene	/codon_start=2				
gene	/product="lecithin-cholesterol acyl transferase"				
gene	/protein_id="AAF22841.1"				
gene	/db_xref="GI:6653738"				
gene	/translation="KFRVYSKFFLDLDKLIENASTSNGRPVILVSHLSGLFALHLILNRPNTSWRRKFIKHEVLSIPWGTVDLMTFASGNLGVLPDLVREORSESNLMLPNPKLGRKTRPLVITPSPASDIPQLSDIFEGQVHFKEKTRVLIGLMQLVAMPGLVAFPMKCEKDSGDFVLMKQXVMSSLRPFMGDDGQVNVMSLMAHSLMSDMKKNQTLKTRIRIGVSTFALIKDEGALDOIIVREISGNSQLGFENNAISVE"				
BASE COUNT	230 a 176 c 194 g 231 t 2 others				
ORIGIN					
Query Match	27.7%	Score 369.4;	DB 8;	Length 833;	
Best Local Similarity	70.5%	Pred. No. 2,1e-87;			
Matches 532;	Conservative 2;	Mismatches 218;	Indels 3;	Gaps 3;	
544 GTGGTTCGAAGTTCCTCAAGATCTTAAGAATTTGATAGAGAAGCAAGCAATTCGAT	603				
11 GTTATTTCCAGTTCCTTCAAGACCTTAAGAAATTTGATAGAAAAGCAAGCACCTCTAAT	70				
604 AATGGGAAGCCAGATACCTCTGCCACAGATTAGAGGCGCTATTTTGTCTACACACTA	663				
71 GGAGGAAGGCCAGTATTTCTGTGTACACAGCTTAGAGAGGCTTTGTCTCCACCTC	130				
664 CTAAATGAAAACCCCTCTTGGGCAAAAAATTGATCAACAACACTTCATGCTCTTCA	723				
131 CTCAACGGAGACACACCTTCGTGGGGGTGGAAATTCATCAACAACCTTTGTTCACTCC	190				
724 GCTTCATGGGGTGTCTATAGAGCAAAATGTACACCTTTGCATCTGGCAACACTTTGGGA	783				

Db	191	ACACATG6GGGTHGACAGTGGACGAAATGCTCACATTTGGCTCAGGCAACACATTTAGG	250
Qy	784	GTGCCCCCTAGTGGACCCCTTTATTAGTAGGAGATGACAAAGAACTCCGAGAGTAACTT	843
Db	251	GTGCCCCCTAGTGGATCCATTTGCTCTAAGAGAGAAAGCAAGAGACTCAGAGAGCACTT	310
Qy	844	TGGCTTTTGGCCTAACCCAAAATTTTGTGGTCTCAAAAACCAATGATGTAATCCCAATT	903
Db	311	TGGCTAATGCCAAACCCCTAATTTATTTGGTGTGTAAAACCCCACTTGATATACCCCAAGC	370
Qy	904	AGGCCCTTATTCAGCTCATGACATGGTTGATATTTTCAAAAAGACATTTGTTCTCTGAAGG	963
Db	371	GCTACATATTCAGCTCGCTCGATATTCACACAATTTCTCAGAGACATTTGATTTGAACAAGT	430
Qy	964	GTTTATTCCTTATGAACAAGAAATCTACCCCTTGATAGGAACATATAAGCACCACA-AGT	1022
Db	431	TTTTCATTCCTTTCAAAACCCGCGCTTTTGGGCTTGATGATCAGCTTGTTCCCGGGGGT	490
Qy	1023	GCCCTAATCTGTATATAGGGAAGCGAGGTGGAACTTGGAACATATGTTTATAGGAA	1082
Db	491	GCCATATTCCTGTGTGTTTGGATGAGTGGTGAAGACAGCTTGAGACTTGTCTATAGAAA	550
Qy	1083	-AGTGATTTTGGATGAACGGCCAGAAATATCATAT-GGGAGATGTGATGAAGAAAGGTGAAC	1140
Db	551	CAARGGTTTTGATGATGACGACGCTGAGATTTGTTATAGGGGATGAGATGGAACCTGTGAAC	610
Qy	1141	TTTGTGACCTTTGTGGCGCTTCATCTACTATGGAAGAAGCAAAAATCAATACCTTAAA	1200
Db	611	ATGGGACCTTAATGGCACATGATFCTCTATGCTGTGATGAAATAACCAACCTTAAAA	670
Qy	1201	GTGGTAAATGATGAGGGGTGCTCATCTTCATCAATCACTTAAGATGAAGTGGCACTAAT	1260
Db	671	ACCATCAAAATTCCTGGGGGTTCTCATACGGCCATATCAAGAAGCAAGGTGCCTTGAT	730
Qy	1261	GAATATGATGAGGATTTACTTCAATTTATTTCTCA	1295
Db	731	CAATTTGTCAAGAAATTTCTGGGATTTAATTCCTCA	785
RESULT 6			
AX037586		1562 bp	DNA
LOCUS	AX037586		Linear
DEFINITION	Sequence 10 from Patent WO0060095.		PAT 16-NOV-2000
ACCESSION	AX037586		
VERSION	AX037586.1	GI:11227005	
KEYWORDS			
SOURCE			
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 1562)		
REFERENCE	Banas,A.; Stahl,U.; Stymme,S.; Lemnan,M.; Ronne,H. and Dahlqvist,A.		
AUTHORS			
TITLE	A new class of enzymes in the biosynthetic pathway for the production of triacylglycerol and recombinant dna molecules encoding these enzymes		
JOURNAL	Patent: WO 0060095-A 10 12-OCT-2000; BASF PLANT SCIENCE GMBH (DE); BANAS ANTONI (PL); STAHL ULF (SE); STYME STEN (SE); LEMMAN MARIT (SE); RONNE HANS (SE); DAHLQVIST ANDERS (SE)		
FEATURES			
source	Location/Qualifiers		
	1..1562		
	/organism="Arabidopsis thaliana"		
	/db_xref="taxon:3702"		
BASE COUNT	441 a 312 c 367 g 442 t		
ORIGIN			
	Query Match 22.4%; Score 298.2; DB 6; Length 1562;		
	Best Local Similarity 61.4%; Pred. No. 1.9e-68;		
	Matches 527; Conservative 0; Mismatches 308; Indels 24; Gaps 2;		
Qy	440	TGATGATTCATTAACAAAGCTTGCTAGCTGATGTGATGACATCTCTGTTGGAGCCCTT	499

Db 694 TGAAGCTCTAGAGAAAAATGCGGTATGTTAAGACCAACCATCTAGAGCTCCAT 753
QY 500 ATGACTTATAGATGTCTAGCTGTAAGGTACACCTTTCAGATGGGTTCCAGTTCC 559
Db 754 ATGATTTTCAGTAGAGGCTGCTGCTGCGGCCACCCGTCCTCCGTAGACCTCACAGTTCC 813
QY 560 TCAAAAGTCTAAGAAATTTGATAGAGAGCAAGCAATTCATTAATGGAAGCCAGTA 619
Db 814 TACAGACCTCAACAAATTTGGTGAAGAAAACCTACAGAGAGAAAGAAAGAGAGTA 873
QY 620 TACTCTCTCCACAGTTTAGAGGCTATTGTCCTACAACTACTAATTAAGAAACCCC 679
Db 874 TACTCTCTCCACAGTTTAGAGGCTATTGTCCTACAACTACTAATTAAGAAACCCC 933
QY 680 CCTCTGCGCAAAAATTAATCAACACTTCTGCTTTCAGCTCCATGGGTGGTG 739
Db 934 CTTCATGCGCGCAAGATACATCAACACTTTCAGCTCCATGGGTGGTG 993
QY 740 CTATAGAGAAATGTACACCTTTCATCTGCAACACTTGGAGTGCCTTAGTGACC 799
Db 994 CGATCTCTAGATAGAGACATTTGCTTCTGCAACACTTGGAGTGCCTTAGTGACC 1053
QY 800 CTTTATTAGTAGAGGATGAGCAAAAGACTCCAGAGTAACCTTTGGCTTTGCTTAAC 859
Db 1054 CTTTCTGCTGAGAGGATGAGCAAAAGACTCCAGAGTAACCTTTGGCTTTGCTTAAC 1113
QY 860 CAAAATTTTGGTCTCA---AAACCAATAGTATGATACCTCAATTAAGGCTTAATTCAG 916
Db 1114 CCAAGTGTTCACAGACAGAACTAACCGCTGCTGTAACCTCCAGAGTTAACTAACAG 1173
QY 917 CTCATGACATGTTGATTTTCTAAGACATGCTTTCCTGAGAGGCTTATCTTAATG 976
Db 1174 CTTACAGATGAGATGAGTTTTCCTGACAGATGATTTCTCAAGAGATTTGGCTTAAC 1233
QY 977 AACACGAATTTACCTTGTATAGGAGAACTAAAGACACCAAGTCCCTAATTAATCTGTA 1036
Db 1234 AGACAAAGTGTGCTTAAACAGAGAGTGTATGATCCGAGAGTCCCACTCACTTGA 1293
QY 1037 TTATGGAGGAGGAGTGGGAAACCTTGAACATGTTTATGGAAGGTGATTTGATG 1096
Db 1294 TATATGGAGAGGAGTGTATGATACCGGAGTGTGATGATGAAGAGAGATTCGATA 1353
QY 1097 AACGCCAGAAATATCATATGAGGAGTGTATGAGAGGATGATGCTGCTTGTG 1156
Db 1354 AGCAGCAGAGATTAATATGAGAGATGAGATGAGAGGATTAATTTGGCGAGCTTAGCAG 1413
QY 1157 CGCTTCAATCACTATGGAAGAGAGAAATCAATACCTTAAAGTGGTTAAGTAGATG 1216
Db 1414 CTTT-----GAAAGTCATAGCTTGAACACCGTAGAGATTTGATG 1452
QY 1217 GGTGTCTCATCTTCAATCTTAAAGATGAGTTCACATTAATGAATAGTAGTGAGA 1276
Db 1453 GAGTTTGCATACATCTATCTTAAAGACGAGATGCACCTTAAAGAGATTAAGAGAGA 1512
QY 1277 TTACTTCATTAATTTCTCA 1295
Db 1513 TTTCATTAATTAATTAATGA 1531

RESULT 7
AX037605 1562 bp DNA linear PAT 16-NOV-2000
LOCUS Sequence 29 from Patent WO0060095.
DEFINITION AX037605
ACCESSION AX037605
VERSION AX037605.1 GI:11227019
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 1562)
REFERENCE
AUTHORS Banas,A., Stahl,U., Szymme,S., Lenman,M., Ronne,H. and Dahlqvist,A.

TITLE A new class of enzymes in the biosynthetic pathway for the production of triacylglycerol and recombinant dna molecules
JOURNAL BASF PLANT SCIENCE GMBH (DE) ; BANAS ANTONI (PL) ; STAHL Ulf (SE) ; SZYMME STEN (SE) ; LENMAN MARIT (SE) ; RONNE HANS (SE) ; DAHLQVIST ANDERS (SE)
FEATURES location/Qualifiers
source 1. 1562
BASE COUNT 441 a 312 c 367 g 442 t
ORIGIN
Query Match 22.4% Score 298.2; DB 6; Length 1562;
Best Local Similarity 61.4%; Pred. No. 1.9e-66;
Matches 527; Conservative 0; Mismatches 308; Indels 24; Gaps 2;
QY 440 TGTAGATTTCAATTAAGAGCTTGGCTGATGATGAGACCTGTTGGAGCCCTT 499
Db 694 TGAAGCTCTAGAGAAAAATGCGGTATGTTAAGACCAACCATCTAGAGCTCCAT 753
QY 500 ATGACTTATAGATGTCTAGCTGTAAGGTACACCTTCAAGTGGTTCCAGTTCC 559
Db 754 ATGATTTCAAGTAGAGGCTGCTGCTGCTGCGGCCACCCGTCCTGATGCTTCAAGTTCC 813
QY 560 TCAAAAGTCTAAGAAATTTGATAGAGAGCAAGCAATTCATTAATGGAAGCCAGTA 619
Db 814 TACAGACCTCAACAAATTTGGTGAAGAAAACCTAGAGAGAGAGAAAGAAAGCCAGTA 873
QY 620 TACTCTCTCCACAGTTTAGAGGCTTATGTCCTCACTACTAATTAAGAAACCCC 679
Db 874 TACTCTCTCCACAGTTTAGAGGCTTATGTCCTCACTACTAATTAAGAAACCCC 933
QY 680 CCTCTGCGCAAAAATTTGATCAAAACCTTATGCTCTTTCAGCTCCATGGGTGGTG 739
Db 934 CTTCATGCGCGCAAGATGATCAAAACCTTGTGTAACCTCCAGAGTATGGGTGGGA 993
QY 740 CTATAGAGAAATGTACACCTTTCATCTGCAACACTTGGAGTGCCTTAGTGACC 799
Db 994 CGATCTCTAGATAGAGCAATTTGCTTGGCAACACCTGCTGCTTATGATTAAC 1053
QY 800 CTTTATTAGTAGAGGATGAGCAAAAGAGTCCGAGAGTAACCTTGTGCTTAAC 859
Db 1054 CTTTCTGCTGAGAGGATGAGAGGATGAGAGACTTTCGCTCATTCTCAGCTTAAC 1113
QY 860 CAAAATTTTGGTCTCA---AAACCAATAGTATGATACCTCAATTAAGGCTTAATTCAG 916
Db 1114 CCAAGTGTTCACAGACAGAACTAAACCGCTGTCGTAACCTCCAGAGTTAACTAACAG 1173
QY 917 CTCATGACATGTTGATTTTCTAAGACATGCTTTCCTGAGAGGCTTATCTTAATG 976
Db 1174 CTTACAGATGAGATGAGTTTTCCTGACAGATGATTTCTCAAGAGATTTGGCTTAAC 1233
QY 977 AACACGAATTTCTACCTTGTATGAGAGGATGATGCTGCTTGTG 1036
Db 1234 AGACAAAGTGTGCTTAAACAGAGAGCTGATGATCCGAGAGTGCATCACTTCA 1293
QY 1037 TTATGGAGAGGAGTGGGAACCTTGAACATGTTTATGGAAGAGTATTTGATG 1096
Db 1294 TATATGGAGAGAGGATGATACACCGAGTGTGATGATGAAGAGAGATTCGATA 1353
QY 1097 AACGCCAGAAATATCATATGAGGAGTGTATGAGAGGATGATGCTGCTTGTG 1156
Db 1354 AGCAGCAGAGATTAATATGAGAGATGAGATGAGAGGATTAATTTGGCGAGCTTAGCAG 1413
QY 1157 CGCTTCAATCACTATGGAAGAGAGAAATCAATACCTTAAAGTGGTTAAGTAGATG 1216
Db 1414 CTTT-----GAAAGTCATAGCTTGAACACCGTAGAGATTTGATG 1452
QY 1217 GGTGTCTCATCTTCAATCTTAAAGATGAGTTCACATTAATGAATAGTAGTGAGA 1276
Db 1453 GAGTTTGCATACATCTATCTTAAAGACGAGATGCACCTTAAAGAGATTTAAGAGAGA 1512

[illegible]

FEATURES Location/Qualifiers
 source 1..264
 /organism="Glycine max"
 /db_xref="taxon:3847"
 unsure 39
 /note="n=unknown"
 unsure 175
 /note="n=unknown"
 241
 /note="n=unknown"
 BASE COUNT 62 a 73 c 52 g 74 t 3 others
 ORIGIN

Query Match 18.7%; Score 249; DB 6; Length 264;
 Best Local Similarity 98.5%; Pred. No. 2.7e-55;
 Matches 260; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

OY 315 CCAAGAACGCGANTACTTCAACACCTCGGTGGGTGAGACCGGTCCTCACCTTTGG 374
 |||
 DB 1 CCAAGAACTCGANTACTTCAACACCTCGGTGGGTGAGACCGGTCCTCACCTTTGG 60

OY 375 TTGCACCAACTCTCTTCT-ATCTCAATCTCTGCTCAAGCATATCACCGATACATGG 433
 |||
 DB 61 TTGCACCAACTCTCTTCTCATCTCAATCTCTGCTCAAGCATATCACCGATACATGG 120

OY 434 CACCCCTGTAGATTCAATCAAAAGCTTGCTAGCTGATGATGAGACTCTGTTTGGAG 493
 |||
 DB 121 CACCCCTGTAGATTCAATCAAAAGCTTGCTAGCTGATGATGAGACTCTGTTTGGAG 180

OY 494 CCCCTTATGACTTTGATGATGCTAGCTGCTGAAAGGTACCCCTTACAAAGTGGTTCCA 553
 |||
 DB 181 CCCCTTATGACTTTGATGATGCTAGCTGCTGAAAGGTACCCCTTACAAAGTGGTTCCA 240

OY 554 AGTTCCTCAAGATCTAAAGAATT 577
 |||
 DB 241 NCTTCCTCAAGATCTAAAGAATT 264

RESULT 10
 LOCUS AX090320 273 bp DNA linear PAT 21-MAR-2001
 DEFINITION Sequence 13 from Patent WO0116308.
 ACCESSION AX090320
 VERSION AX090320.1 GI:13444186
 KEYWORDS
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 REFERENCE 1 (bases 1 to 273)
 AUTHORS Lasserre M. and van Eenennaam A.
 TITLE Plant sterol acyltransferases
 JOURNAL Patent: WO 0116308-A 13 08-MAR-2001;
 MONSANTO COMPANY (US)
 FEATURES
 source Location/Qualifiers
 1..273
 /organism="Glycine max"
 /db_xref="taxon:3847"
 12
 /note="n=unknown"
 33
 /note="n=unknown"
 252
 /note="n=unknown"
 265..266
 /note="n=unknown"
 272
 /note="n=unknown"
 BASE COUNT 80 a 56 c 62 g 69 t 6 others
 ORIGIN

Query Match 15.5%; Score 206; DB 6; Length 273;
 Best Local Similarity 97.7%; Pred. No. 7.6e-44;
 Matches 217; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

OY 1 ATGAGAAAGAACCAAGAGGGTCTCAAGATTGAGGTCTTACACTCAGCTTACAGTA 60
 |||
 DB 53 ATGAGAAAGAACCAAGAGGGTCTCAAGATTGAGGTCTTACACTCAGCTTACAGTA 112

OY 61 GTTGTGTGATGCTGTCTATGTCATGATGAGGGGCAAGCAACCTGACCTTTGATT 120
 |||
 DB 113 GTTGTGTGATGCTGTCTATGTCATGATGAGGGGCAAGCAACCTGACCTTTGATT 172

OY 121 CTAATACCAAGTACGAGGAGCAACCACTAGAACAGGTTACCAATCAGTACAGCCC 180
 |||
 DB 173 CTAATACCAAGTACGAGGAGCAACCACTAGAACAGGTTACCAATCAGTACAGCCC 232

OY 181 TCTACTTTCATCTCGAATCATGATACCTCTCATCAGAAA 222
 |||
 DB 233 TCTACTTTCATCTCGG-ATCNTGTGATCCTCTCANNAAGANA 273

RESULT 11
 AC120983/C
 LOCUS AC120983 134188 bp DNA linear HTG 21-JUN-2002
 DEFINITION Oryza sativa chromosome 3 clone OSJNB0011H13, *** SEQUENCING IN PROGRESS ***, 4 ordered pieces.
 ACCESSION AC120983
 VERSION AC120983.1 GI:20564440
 KEYWORDS HTG; HTGS; PHASE2.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
 REFERENCE 1 (bases 1 to 134188)
 AUTHORS Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Tallon, L., Gansberger, K., Kim, M., Overton, II, L., Bera, J., Tsilini, T., Krol, M., Jarrahl, B., Jin, S., Koo, H., Zismann, V., Hsiao, J., Blunt, S., Vanaken, S., Uteerbach, T., Feldblyum, T., Yang, Q., Haas, B., Suh, B., Peterson, J., Quackenbush, J., White, O., Salzberg, S. and Fraser, C.
 Oryza sativa ssp. japonica cv. Nipponbare OSJNB0011H13 BAC genomic sequence
 TITLES
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 134188)
 AUTHORS Buell, R.
 JOURNAL Direct Submission
 COMMENT Submitted (14-MAY-2002) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
 * NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces * is believed to be correct as given, however the sizes * of the gaps between them are based on estimates that have * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and * the accession number will be preserved.
 * 1 61218: contig of 61218 bp in length
 * 61219 61256: gap of unknown length
 * 61257 77652: contig of 16396 bp in length
 * 77653 77727: gap of unknown length
 * 77728 131160: contig of 53433 bp in length
 * 131161 131197: gap of unknown length
 * 131198 134188: contig of 2991 bp in length.
 FEATURES
 source Location/Qualifiers
 1..134188
 /organism="Oryza sativa"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /chromosome="3"
 /clone="OSJNB0011H13"
 /note="japonica cultivar-group"
 BASE COUNT 39144 a 28013 c 28230 g 38642 t 159 others

ORIGIN

Query Match	14.8%;	Score 197.6;	DB 2;	Length 134188;
Best Local Similarity	50.9%;	Pred. No. 9.7e-42;		
Matches 650;	Conservative 0;	Mismatches 534;	Indels 92;	Gaps 4;

OY	104	ACGCGACCCCTTGAATCTATATACAGGATACGGGAGGGAACCAACTATACAGACGGTTGA	163
Db	77046	AGCTGCACCCCGGTGATCTGTATCCCGGCGCGCGGGGAACCAAGTGAAGGCGCGGTGA	76987
OY	164	CCAATCAGTACAAAGCCCTCTACTTTCATCTCATCTCGAATCATGAGACCTCTCATCAAGAAAA	223
Db	76986	CGAGGAGATACGCCCCGCTCGAGCCCTCGGGTGTCCGGGTGTGG---CCCGTGGTGGCGGCG	76930
OY	224	AGAAATGATGCTTCAGACTTTCGTTTGATTCACAGTGTCAATCTGCTGCTCTTACTCAAT	283
Db	76929	GGCGGGGTGGTTCGCGCTCTGTGTTCCAGCCCGTCCGTGTGTGGCGCGGCTCACAGAGT	76870
OY	284	GCTTTGGCGAAGCGATACCCCTCATTTTCCACCAAGAACTGGATATTAATCTTCAACACTC	343
Db	76869	GCTTGGCCGACCGGATGACCCCTCTTTACAGATCCGTCGCCGACGATCAACGCAAGCTC	76810
OY	344	CTGGGGTTGAGACCCCGGGTCCCTCACTTGGTTCCACCACACTCTCTTTCTATCTCAATC	403
Db	76809	CCGGGCTGAGAACCAAGGGTCTCCGACTTCGGGTCCACTCCACCTCCGCTGACCTCGAC	76750
OY	404	CTCGTCTCAACATC-----	418
Db	76749	CCAACCTCAAGTACGATTTCTACTATCAGCATTCACAGCGTGTGTTGGTGTAT	76690
OY	419	-----TCAACCGGATCATGGACCCCT	440
Db	76689	GACCGCAATGCAAAATGCAAAATGGGATGTGTAGTTGCTTAACGGGGTACATGAACGTGTT	76630
OY	441	GGTAGATTCATTACAAAAGCTTGGCTGACGTGATGGTGAAGACTGTTTGAAGCCCTTA	500
Db	76629	GGCGAGCAGCGCTGGAGAGGACGAGGTACGAGGAAGGTTGCACTGTTCCGGCGGCGTA	76570
OY	501	TGACTTTTGATATGTGTCTAGCTGCTGAAGGTACCCCTTACAAAGTGGGTTCCAACTTCT	560
Db	76569	CGACTTCGGTGAAGGGCTGGCGGGGCGGGGACACCGCTGGGTGTGGGACAGGCGCTACT	76510
OY	561	CAAAAGATTAAGAAATTTGATAGAAAGCAAGCAAGCAATTCATAATGGGAAGACAGTAT	620
Db	76509	GGAGCGGCTGAGGAAGCTGTGTGAAGTCCGGCTGCGCGGCCCAACGGCGGGAGCGCGGAT	76450
OY	621	ACTTCTCTCCACAGATTAGAGGCGCTATTGTCTCTACACTACTAATAATAGAACCCGCC	680
Db	76449	CTCTGTGGCGACACCGCTGGCGGGGCTGTACGCGCTGACAGATGGTGGCGGAGGCCGCC	76390
OY	681	CTCTTGGCGCAAAAATTCATCAAAACACTTCATTCCTTTTCAAGTCCATGGGGTGTGTC	740
Db	76389	CGCGGGGCGGCCCGGGAAGGTGAAGCGGGCTGGTGAAGCTGTCCGGCGCGTGGGGGGGCTC	76330
OY	741	TATAGACCAAAATGTACACTTTGCTATCTGGCAACACTTTGGGAGTGGCCCTAGTGAGCC	800
Db	76329	GGTGAAGAGATGTCTACACTTCGCGCTCGGGCAACACCTCGCGGTGCCGTTGTGTGAGCG	76270
OY	801	TTTATTTAGTGAAGGTGTAACAAAGAGCTCCGAGAGTAACTTGGCTTTTGTCTAACCC	860
Db	76269	GTCGCTCATCCGCGACGAGCGACCGCACCGCGAGAGCAACTGTGTGGTGTGCCCGGCC	76210
OY	861	AAAAATTTTGGTCTCAAAAACCAATGT---GATAACTCCAATTAGCGCTTATTCAGC	917
Db	76209	GGGGTGTTCGGGAACACCAACGCTGTGTGTGTCGGGCAACCAACCGGCTCTACTCCGC	76150
OY	918	TCATGACATGTTGATTTTCTAAAAGCAATTTGTTTCTGAAAGGGTTTATCTTATGA	977
Db	76149	CAAGAATCATGACGAGCTTCCTCCGGGACATCGGGTTTCCCGAAGGGGTGGAGCGCTACCG	76090
OY	978	AACACGAATTTACCTTGTGATAGGGAACATAAAGCAACACCAAGTGCCTATTAACTTTAT	1037
Db	76089	GGAGGGATTTAGCGCGCTGTCTGAGAGTCTGCGCGGAGCCCGGGGTGCCGGTATACCTGCT	76030

QY	1038	TATGGAGACGGAGTGGGACCTTGGAAACATTGTTTATGGGAAAGGATTTGATGA	1097
Db	76029	CGTCGGACAGGGGCGTCGACACCTGGAGAGCTCGTGTACGGGAGGGGGCTTGAGGC	75970
QY	1098	ACGGCCAGAAA---TATCATATGGGATGTGATGGAACGGGAACCTTGAGCTGTT	1154
Db	75969	CGGGCGCGGAGGAGGTGTGTACGGGACGGGACGGGATCAACCTTCGACAGCTCGT	75910
QY	1155	GGCGCTTCAATCATATGGAAGAGGAGAAATCAATACCTTAAAGTGGTTAAACATAGA	1214
Db	75909	CGGGCGCATCAAGCGGTGTCGACTCGCGGAGGAGGTGGTGGAGGTGTCGACTGCC	75850
QY	1215	TGGGGTGTCTCATCTTAATCTTAAGGATGAATGGTGCATTAATGAATAGTAGTGCA	1274
Db	75849	CGAGGTGTGCGACCTCCGGATCTCTCAAGGACAAGAGCGGCGTCGACCAATCTTGAGGAT	75790
QY	1275	GATTACTTCATTAAT	1290
Db	75789	CATGACGCGCATCAAT	75774

RESULT 12	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORANISM
AC103891	123620 bp	DNA	linear	PLN 26-APR-2002			
Oryza sativa	(japonica cultivar-group)	chromosome 3	clone				
Oj175C11	complete sequence.						
AC103891							
AC103891.2	GI:20279375						
HTG.							
Oryza sativa	(japonica cultivar-group).						
Oryza sativa	(japonica cultivar-group)						

Eukaryota: Viridiplantae; Streptophyta: Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 123620)

AUTHORS	Wing, R.A., Yu, Y., Soderlund, C., Chen, M., Kim, H.-R., Rambo, T., Sasaki, C., Henry, D., Oates, R. and Simmons, J.
TITLE	Rice Genomic Sequence
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 123620)
AUTHORS	Wing, R.A., Yu, Y., Soderlund, C., Chen, M., Kim, H.-R., Rambo, T., Sasaki, C., Henry, D., Oates, R. and Simmons, J.
TITLE	Direct Submission
JOURNAL	Submitted (30-NOV-2001) Clemson University Genomics Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA
REFERENCE	3 (bases 1 to 123620)
AUTHORS	Wing, R.A., Yu, Y., Soderlund, C., Chen, M., Kim, H.-R., Rambo, T., Sasaki, C., Henry, D., Oates, R. and Simmons, J.

REFERENCE	TITLE	AUTHORS
4 (bases 1 to 123620)	Direct Submission Submitted (24-APR-2002) Clemson University Genomics Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA	Wing, R.A., Yu, Y., Soderlund, C., Chen, M., Kim, H., Rambo, T., Saski, C., Henry, D., Oates, R., Stimmans, J., Thurmond, S. K. and Sun, S.
	Direct Submission Submitted (26-APR-2002) Clemson University Genomics Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA	
	On Apr 24, 2002 this sequence version replaced gi:117155018.	
	This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >=30). An attempt was made to resolve all sequencing problems, such as compressions and repeats: all regions were covered by more than one plasmid subclone; and the assembly was confirmed by restriction digest. Bacterial transposons can be found throughout the project at these locations: 120370-123403, 96027-98926, 96146-98910, 94511-97689, 64299-66588 and 11335-14899. From base 43413-43437, there is only a single subclone. There are 2 bases below standard at 96468-69. At base 93692 there is either a G or a pad - Clemson University reads disagree with Monsanto data here. The nucleotide sequence of this BAC clone was generated by combining Monsanto and Clemson University Genomics Institute	

FEATURES		sequencing data.
source	Location/Qualifiers	
gene	1..123620	
	/organism="Oryza sativa (japonica cultivar-group)"	
	/db_xref="taxon:39947"	
	/chromosome="3"	
	/clone="OJ1175C11"	
CDS	complement(4328..5515)	
	/gene="OJ1175C11.1"	
	/note="Contains similarity to peroxidase"	
	complement(join(4328..4921,5019..5210,5297..5515))	
	/gene="OJ1175C11.1"	
gene	/codon_start=1	
	/product="Putative peroxidase"	
	/protein_id="AA19121.1"	
	/db_xref="GI:20330758"	
	/translation="MAAASVVICAVLVLAAGASAPVMSPSYFASGCPYVDIY RRVVEARCTPPRAPASLRLRHFDICFVNGCDGSLLDDEGAMOSEKAPPKGSARG FDVVDGIAALEMNCVGVSCADILALAEISVELISGSPSNVMIGRBDGVAANEGA RDLPPTDDLDLRLRRKFSFENLDIDTDFALOGAHTIGRAOCRFHDLRLYNTSGTEPD QTLDMAYLNEIROSCPSADPESADRLNDPPTPDADFNDSYGNLLRNGLLOSOGML SAPGGAASTAPIVVMFAGSQDDFFRSFATAMVKGNISPLTSGMGEIRNRQVNRG"	
gene	8150..10413	
	/gene="OJ1175C11.2"	
	/note="Contains similarity to ubiquitin / ribosomal	
	protein CEP52"	
	join(8150..8252,8344..8533,8815..8940,10042..10108, 10384..10413)	
CDS	/gene="OJ1175C11.2"	
	/codon_start=1	
	/product="Putative ubiquitin / ribosomal protein CEP52"	
	/protein_id="AA19122.1"	
	/db_xref="GI:20330759"	
gene	/translation="MQIFVKLTGKTTILEVSSPTIDNVAKIODEKCIPTDQRLI FAGQLEDGRTLADINIKESTLHLVRLRGITLIPSLQALARKTQNRKNCRKD SLRSLPSGPEAFSEEMQDKLWGLPDLMLRLPVDKRFSCVRLHRAVNRCKKKGHS NQLRRKKIKIN"	
	complement(11559..12591)	
	/gene="OJ1175C11.3"	
	/note="Putative hemoglobin 4"	
	complement(join(11559..11705,12033..12149,12260..12374, 12467..12591))	
CDS	/gene="OJ1175C11.3"	
	/codon_start=1	
	/product="Putative hemoglobin 4"	
	/protein_id="AA19123.1"	
	/db_xref="GI:20330760"	
gene	/translation="MAFASNSGAVRTEQEDALVLSMAITMKDSDSANIGHRFLKIF EVAPSARHLFSFLRNSDVLKPNPLKKHMAVFMTCESAAQLKRTGRVVRDTTIK RLGSTHFKNVSDTHFEVAPRALLETIKDGIIPASWMSPEMKANCGEAYEHLVAAIKEG MKPVALL"	
	complement(15258..16098)	
	/gene="OJ1175C11.4"	
	/note="Putative hemoglobin 3"	
	complement(join(15258..15401,15490..15606,15700..15814, 15965..16098))	
CDS	/gene="OJ1175C11.4"	
	/codon_start=1	
	/product="Putative hemoglobin 3"	
	/protein_id="AA19124.1"	
	/db_xref="GI:20330761"	
gene	/translation="MAANGSNVSRGAVRTEQEDALVLSMAITMKDSDSANIGHRFL KIFEVAPSARHLFSFLRNSDVLKPNPLKTHAMAVFMTCESAAQLKRTGRVVRDTTIK TRLGSTHFKNVSDTHFEVAPRALLETIKDGIIPASWMSPEMKANCGEAYEHLVAAI KQMKPFAA"	
	complement(18978..19804)	
	/gene="OJ1175C11.5"	
	/note="Putative Non-symbiotic hemoglobin 1"	
	complement(join(18978..19118,19234..19350,19446..19560, 19677..19804))	
CDS	/gene="OJ1175C11.5"	
	/codon_start=1	
repeat_region	repeat_region	
	repeat_region	
	repeat_region	
	repeat_region	
	repeat_region	
gene	/product="Putative Non-symbiotic hemoglobin 1"	
	/protein_id="AA19125.1"	
	/db_xref="GI:20330762"	
	/translation="MAIVEDNNAVASFSEEOALVLSMAITMKDSDSANILREFIKI FEVAPSASOMCSFLRNSDVLKPNPLKTHAMAVFMTCESAAQLKRTGRVVRDTTIK RLGATHLKTVGVDHFEVAPFALDITKEEYPADMSPAKSAWSEAYDHLVAAIKO EMKPAE"	
	complement(20266..20413)	
CDS	/note="Similar to Tourist O11 MTE element"	
	complement(20716..23919)	
	/gene="OJ1175C11.6"	
	/note="Hypothetical protein"	
	complement(join(20716..21204,21467..21553,21790..21889, 21965..22371,22488..22925,23037..23153,23496..23579, 23869..23919))	
repeat_region	/gene="OJ1175C11.6"	
	/codon_start=1	
	/product="Hypothetical protein"	
	/protein_id="AA19126.1"	
	/db_xref="GI:20330763"	
CDS	/translation="MDGTDLSPPRCNVQHLQNAEELKDONSTNRKLPPTTEPCSLIO EVOHEKRLNDOPAMRALKEKALCYKPCALIHSSNSCIPKPTTELIKETIAVLELVIC LEOHLAIYKRAFPDOICVSSSCSMEINKSASFSITLGSSELDSTYRKQILLO SSGVMARKSTPTLTSETRTSHNDKGTIGRSISLQRSICSAVSPSANNLALAL KPCHTLPLSFVEEGKMDPGIVSLADILGTIRIADHPQTPRKISEDMIKTASITIRI RDFNAVQHPFPSPCSFSASGSKTGIDWIPCRKEGIEIAMODALGTESRY FSOOYDSVIEVSALCGKORSADVKDMHKYKSLVOLLESADLGMKMEKTIAPFNIY HNAMMHMLSTYLISGQRVNDELIEYHITICCRHSPTOMRLILYKMSKEKEDOGFA VDREPLVHFALSSGSHSDPYVRLKRPRLLOQIEARDEVEVLRVAVNGYCGRCGRGR VLLPLKLEPYSRDAGLGHDLRAVSCLEPELRPAQQAARSRGGGGVEMRPHN PAFRYLARELVGPPAPVHLST"	
	20811..20828	
	/note="Simple Sequence Repeat (CGC)n"	
	21305..21425	
	/note="Similar to SINEIOS Rice p-SINEI retroposon"	
CDS	28171..30050	
	/gene="OJ1175C11.7"	
	/note="Hypothetical protein"	
	join(28171..28347,28968..29037,29138..29228,29490..29844, 29961..30050)	
	/gene="OJ1175C11.7"	
repeat_region	/codon_start=1	
	/product="Hypothetical protein"	
	/protein_id="AA19127.1"	
	/db_xref="GI:20330764"	
	/translation="MAAANLHAHPPPVPLNARAVGAPPEKRIEVRERETLSPGR AAAHAAVAFKVENQVEYNDLTCRQRTFGQWNAQOAGANENNSGCDDEPPTQ ODVPLETIGITMLHNNLRAITSEYKSIDMETEKAMADKKRIADVDYEGNRKYN TSSERKACVDEGKSIDAEITDDEGVGEKAKCIAERKSISDAETDDEGVGVPRLNH KDLHDIMPTYTODYHYKCTNYSVPYCOTSTYEVMEWP"	
CDS	complement(30482..33215)	
	/gene="OJ1175C11.8"	
	/note="Contains similarity to transposase"	
	complement(join(30482..31066,31149..31870,31963..32043, 32139..32213,32948..33215))	
	/gene="OJ1175C11.8"	
repeat_region	/codon_start=1	
	/product="Putative transposase"	
	/protein_id="AA19128.1"	
	/db_xref="GI:20330765"	
	/translation="MEEEYEIPDLNIDPGVOEVLQDEGDGIPDLNLVPAVGEDAFQ YEDELDPNOCFGAHDEHPCGMAVELSNGRSAGEICHLNMPANVGEDIVFDD EYFVLQDSNVADEDEGVYSPYMDVYEDDLHTSSDEEDQONIKRKYLPAAEK AIVGAILASTINGLADRDTEIITAAEDYRRVODIMWRKAKCIAAGVAVDNRKSK PYLNGEAKRSDDLDQVLAALPLNSGCSIRTSASALMAKSOYHRLRYKGAALRHSNJK PILKBAANKQRLFEVCVSLDAATLNAERFDIMONIVIHIDKFNATMKNKTYFLVN EDEPVRQVONKNAIDKYWFLSTLAKPRYDEGNCYFEGKISIMPEVKEPQRSNRN PKGTLYTKSIMVSRSTRAFLITRVPAISACMPREDVGKTIWIOQNALTHILPND AFALAVAGLDITIMNOPNSPDMNVLDGFFASLOSQTYLKSNMDELISVDEE VNEYNNSNLSRVFLNLQACFIEVWKAQGNKYKIPHNKKDLERLDMPLPSLQDIAL YKVMQTLIN"	
	complement(30482..33215)	
	/note="Contains similarity to transposase"	

gene	complement(34294. .37827)
/gene="J01175C11.9"	
/note="Contains similarity to RING zinc finger protein"	
complement(j01n(34294. .34436,34580. .34704,34769. .35317	
36029. .36509,37120. .37396,37580. .37827))	

QY	776	CTTTGGGATGGCCCCCTAGAGGACCTTTATTTATGTAGGGATGTGAACAACCACTCCGACA	835
Db	96379	TTCTCTGGCTCCCAAGATGCTACGAGACTGTCTCCCTGCAACCAATTGTGGAGACTTTTGAAG	96438
QY	836	GTAACTTTGGCTTTTGGCTTAACCCAAAAATTTTGGTCTCAGAAAAACCAATAGTATTA	895
Db	96439	CTAGCATTTGGCCAACTTTCCATCTCCCGGGGCTTTGGG--CAGCAGCCGATCGTGTATCA	96495
QY	896	CTCCAAATTAAGGCGTTATTCAGCTCATGTGACATGCGTTGATTTTCTAANAAGCATTTGGTTTC	955
Db	96496	CAAACCAAGAAACTACTCTGCGCTACGATCTGGAAAGATCTCTTGCCGCGCTTGGTTTGG	96555
QY	956	CTGAAGGGGTTTATCTTATGAACACGAATTTCTACCTTTATGAGGAACATAAAAAGCAC	1015
Db	96556	GCGACGGCATGAGCGCTTTTAGAAGACGAATGGTTGCGAGATGAGCTACTTCGAGGCC	96615
QY	1016	CACAAGTCCCTATACTGTATTTATGAGCAACGGGAGTGGGAACCTTTGGAAACATTGTTTT	1075
Db	96616	CCATGTGTCCTGTCAGCTGCATCAATGGAATGGGAATAAGACGCCACGGCAATTTGGTGT	96675
QY	1076	ATGGGAAGGATGATTTTATGATGAAACGCCCAAGAAATATCATATGAGGGATGGTGAATGAACGG	1135
Db	96676	ACTGGATGGCAACTTCGATGAAACCGGCACAAATGGTGTATGGTATGAGATGGGGCGTG	96735
QY	1136	TGAATTTGGTAGCTTTGTTGGCGCT-----TCAATCACTATGAAAGAGAGAAAAATC	1189
Db	96736	TTTAATCTGATTTAGCTATGTTTGAAGGTTCAACGAGAAATCGCTGCGCAGCCAGGACAGAGGG	96795
QY	1190	AATACCTTAAAGTGGTTTAAGATAGATGGGGGTGCTCACTCACTCAATTAAGATGAAG	1249
Db	96796	GGCATTTCAAGTCGATCAAGGTGCAAAACGCTTCCCATTAAGGGGATTTCTTACGAGTGAAT	96855
QY	1250	TTTGACCTAAATGAATAGTAGGTGAGATTACTTCAATTAATTCATGCTT	1299
Db	96856	GGGGCTTAAAGAGGGTCAATGCAAGGATTTCTTGAAGCGAATTCGGGATTC	96905

RESULT	14
AP004073	
LOCUS	AP004073 147548 bp DNA linear PLN 13-MAR-2002
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC clone:P0614D08.
ACCESSION	AP004073
VERSION	AP004073.2 GI:19386859
KEYWORDS	
SOURCE	.
ORGANISM	Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA, clone:P0614D08. Oryza sativa (japonica cultivar-group) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Emnathioideae; Oryzeae; Oryza.
REFERENCE	1
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE	Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC clone:P0614D08 Published Only in Database (2001)
JOURNAL	2 (bases 1 to 147548)
REFERENCE	Sasaki,T., Matsumoto,T. and Yamamoto,K.
AUTHORS	Direct Submission
TITLE	Submitted (15-AUG-2001) Takui Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program, Kanonndai 2-1-2, Tsukuba, Ibaraki 305-6602, Japan (Email:sasakientis.affrc.go.jp URL:http://rpg.dna.affrc.go.jp/, Tel:81-258-38-7441, Fax:81-258-38-7468)
JOURNAL	On Mar 12, 2002 this sequence version replaced gi:15208441.
COMMENT	

Genes were predicted from the integrated results of the following databases: GENSCAN 1.0, BLASTX2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI Nonredundant protein database, nr (<http://ncbi.nlm.nih.gov/blast/db>) and the CDNA sequence database at RCGP. Protein homologies of the coding regions were searched against NCBI Nonredundant protein database with BLASTX2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the

corresponding DDBJ accession no. and RCP clone ID. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein. The orientation of the sequence is from Sp6 to T7 of the PAC clone. This sequence of p0614D08 clone has an overlap with P0660F12 (DDBJ:AF003687) clone at the position 1 to 135,297 of 5' end. The sequence of this clone starts at the position 11,907 of P0660F12. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://tgp.dna.affrc.go.jp/GenomeSeq.html>.
Location/Qualifiers

```

source
1..147548
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="P0614D08"
complement(3037..3535)
/gene="P0614D08.1"
complement(3037..3535)
/gene="P0614D08.1"
/note="probably inactive due to frameshift(s) in CDS pseudogene"
/pseudo
complement(join(7123..7412,7452..7583,8141..8285,8890..9097,10827..11083))
/gene="P0614D08.2"
complement(join(7123..7412,7452..7583,8141..8285,8890..9097,10827..11083))
/gene="P0614D08.2"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB86237.1"
/db_xref="GI:19386860"
/translation="MRGAMLPPTTGDMKKTINDGALGCGGGSGGSRITRAGAAARRGESDVAEGRSSILMTAHNRSGGRMTPAVYANWVRGAATSRQPSADSGDRADGSAARSAVVKTEERRGILGVGRKRREEREMSRATHVGEEMRAPWEDSQAAMRRELGDAAATMLVDEADPERLMTTSADDEGVARTAMTTTNGNITVTHGGDTVKKQVHREKKKODRVSRVRCGASDPRERAOAVETKGVGGVGRPMRDGYRSTVKIAROREPDGACGPRRGFEETDEPARLSSGSGGGMRLGIRRSMTVRLGIRSGPRRRRSGGAGAEFDAVAKITRST"
complement(join(11424..11791,11799..12120))
/gene="P0614D08.3"
complement(join(11424..11791,11799..12120))
/gene="P0614D08.3"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB86238.1"
/db_xref="GI:19386861"
/translation="MDWVKRHHGSHGKANGEEEPAGSDCSQRNNAEERRAGGAGARRKREGLKAVSTGSDGEAARMRCPCGGMGVARTPARPAPAASRGSGISLKGARRYEREGRRWRMAERGAAGVASRGREKRRKEEGEADGSGGDDPAVTRDVARREDSRREYTRRRGGGRRSRRRFLAPAVAKKRCCELTDERNHGERTPRRGALMAPRGKDS TAAASALAGK"
complement(join(12756..12787,13386..13509,14014..14210,14640..14763,15312..15440,19176..19286))
/gene="P0614D08.4"
complement(join(12756..12787,13386..13509,14014..14210,14640..14763,15312..15440,19176..19286))
/gene="P0614D08.4"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB86239.1"
/db_xref="GI:19386862"
/translation="MRNNSSEGDSDQIILEMEIYIYEDSELKKKKNGKEIONTPPAAEDSALSVTLWMFSESKNDEDEIEFYIKQGININWTKIIRHTYKTKTSQSDITTRFGGPGSPGPKIESSQSGSPGPNPTDSDDEANTNNSSSGTNGLPNTGTGNR ELGNNLNIHFSTSPSELIELIYIRARPMSPGSGTSDLGISRGISKANTGILPY

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 11:08:52 ; Search time 224 Seconds
(without alignments)
13391.347 Million cell updates/sec

Title: US-09-857-612a-13
Perfect score: 1332
Sequence: 1 atgaagaagaacagaaga.....gtaattgttcggggttaa 1332

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB	ID
1	1332	100.0	1332 21 AAA49206
2	1148	86.2	1217 21 AAA49202
3	449.6	33.8	1299 22 AAS01081
4	449.6	33.8	1495 21 AAC47273
5	448	33.6	1498 21 AAC39845
6	298.2	22.4	1562 21 AAC64437
7	298.2	22.4	1562 21 AAC64447
8	249	18.7	264 22 AAS01087
9	208	15.6	1500 21 AAA49204

10	206	15.5	273	22	AAS01088	Soybean sterol acy
11	144	10.8	563	21	AAA69662	Pinus radiata leci
12	127	9.5	353	21	AAA69561	Pinus radiata leci
13	105	7.9	542	21	AAA49200	Corn putative leci
14	95.4	7.2	1660	21	AAA49205	Corn putative leci
15	86.4	6.5	921	21	AAA49201	Corn putative leci
16	86	6.5	1440	21	AAA49203	Corn putative leci
17	76.8	5.8	1155	22	AAS01095	Corn sterol acyltr
18	64.6	4.8	1323	24	ABK87900	CDNA encoding huma
19	64.6	4.8	1358	8	AAAT70191	DNA clone p12 enc
20	64.6	4.8	1744	20	AA232180	Human lecithin cho
21	64.6	4.8	1744	22	ABAD02648	Human lecithin-cho
22	64.6	4.8	1744	24	ABN95747	Gene #2245 used to
23	64.6	4.8	1744	24	ABK35510	Human endometrial
24	64.6	4.8	4308	22	AAAD02649	DN305 plasmid for
25	64	4.8	1137	19	AAV68563	Murine lecithin-ch
26	64	4.8	1236	19	AAV68568	Murine kidney leci
27	61.2	4.6	4024	23	ABL21222	Drosophila melanog
28	60.4	4.5	1137	19	AAV68561	Human heart leci
29	60.4	4.5	1146	19	AAV68564	Human heart leci
30	60.4	4.5	1236	19	AAV68566	Human heart leci
31	60.4	4.5	2680	20	AA234023	Human PRO540 nucle
32	60.4	4.5	2680	21	AAC78492	Human PRO540 (UNQ3
33	60.4	4.5	2680	21	AAA88516	Human PRO540 CDNA
34	60.4	4.5	2680	22	AAC31556	Human PRO540 CDNA
35	60.4	4.5	2680	24	ABK69967	CDNA encoding huma
36	60.4	4.5	2681	24	AAH14618	Human cDNA sequenc
37	60.4	4.5	2720	22	AAK94279	Human full-length
38	56.8	4.3	1824	23	ABL21223	Drosophila melanog
39	51.2	3.8	1233	19	AAV68562	Human kidney leci
40	51.2	3.8	1242	19	AAV68565	Human kidney leci
41	51.2	3.8	1332	19	AAV68567	Human kidney leci
42	51.2	3.8	2687	21	AAA26392	Human secreted pro
43	47.8	3.6	1941	23	AA591800	DNA encoding novel
44	47.8	3.6	5700	24	ABK87899	Human lecithin-cho
45	47.8	3.6	6901	18	AAAT78853	Human lecithin-cho

ALIGNMENTS

RESULT 1	
AAA49206	
ID	AAA49206 standard; CDNA: 1332 BP.
AC	AAA49206;
DT	12-DEC-2000 (first entry)
XX	Soybean putative lecithin:cholesterol acyltransferase gene #2.
XX	Soybean; lecithin:cholesterol acyltransferase; phytosterol;
KW	phosphatidylcholine-sterol O-acetyltransferase; heat shock; cold shock; ss.
XX	
OS	Glycine max.
XX	
EH	Key
FT	Location/Qualifiers
FT	1..1332
FT	/product= "lecithin:cholesterol acyltransferase"
FT	/EC_number= "2.3.1.43"
XX	
PN	W0200032791-A2.
PD	08-JUN-2000.
XX	
PF	02-DEC-1999; 99WO-us28586.
XX	
PR	03-DEC-1998; 98US-0110782.
XX	
PA	(DUP0) DU PONT DE NEMOURS & CO E I.
XX	
PI	Cahoon RE, Kinney AJ, Sakai H, Shen JB, Butler KH, Saylor JJ;

XX WPI: 2000-412337/35.
 DR P-PSDB: AAB01211.
 XX
 PT Polynucleotide encoding plant lecithin:cholesterol acyltransferase
 PT enzyme useful for producing transgenic plants and for producing
 PT antibodies specific to which is useful for screening cDNA expression
 PT libraries
 XX
 PS
 XX
 PS
 XX
 CC Claim 2: Page 46; 49pp; English.
 CC
 CC The present sequence is a putative coding sequence for a soybean
 CC lecithin:cholesterol acyltransferase (also known as
 CC phosphatidylcholine-sterol O-acyltransferase). This enzyme is found
 CC associated with high-density lipoproteins and Apolipoprotein-AI and -D.
 CC The gene and protein can be used to produce transgenic plants which have
 CC increased lipid metabolism and membrane fluidity, and therefore increased
 CC resistance to heat and/or cold shock, to alter the content of phytosterol
 CC or lecithin in grains and to identify potential herbicides.
 CC
 XX
 XX
 SQ Sequence 1332 BP; 385 A; 282 C; 289 G; 376 T; 0 other;
 Query Match 100.0%; Score 1332; DB 21: Length 1332;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1332: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 721 TCAGCTCATTGGGGGTGTGTATAGACGAATGTACACCTTTGCATCTGGCAACACTTTG 780
 DB 721 TCAGCTCATTGGGGGTGTGTATAGACGAATGTACACCTTTGCATCTGGCAACACTTTG 780
 QY 781 GAGTGTCCCTTACTGACCTTTATTAGTAGGAGATGAACAAAGAGCTCCGAGTAAC 840
 DB 781 GAGTGTCCCTTACTGACCTTTATTAGTAGGAGATGAACAAAGAGCTCCGAGTAAC 840
 QY 841 CTTTGGCTTTTGGCTTAACCAAAATTTTGGTCCCAAAACCAATAGATAGTATCCCA 900
 DB 841 CTTTGGCTTTTGGCTTAACCAAAATTTTGGTCCCAAAACCAATAGATAGTATCCCA 900
 QY 901 ATTAGCCTTTATTACAGCTATGACATGTTGATTTTCTAAACACATTTGTTTCTGAA 960
 DB 901 ATTAGCCTTTATTACAGCTATGACATGTTGATTTTCTAAACACATTTGTTTCTGAA 960
 QY 961 GGGGTTTATCTTATTAACACGAAATTCACCTTATAGGGAACATTAAGCACCACAA 1020
 DB 961 GGGGTTTATCTTATTAACACGAAATTCACCTTATAGGGAACATTAAGCACCACAA 1020
 QY 1021 GTGCTTATACCTGTATTATGGGAACGGAGTGGGAACCTTGGAAACATTTTATGGG 1080
 DB 1021 GTGCTTATACCTGTATTATGGGAACGGAGTGGGAACCTTGGAAACATTTTATGGG 1080
 QY 1081 AAAGTGATTTTGTATGAACGGCCAGAAATATCATATGGGATGTGATGAAACGGTGAAC 1140
 DB 1081 AAAGTGATTTTGTATGAACGGCCAGAAATATCATATGGGATGTGATGAAACGGTGAAC 1140
 QY 1141 TTGCTAGCTTTTGGCGCTTCAATCATATGGAAGAGAGAGAAATCAATACCTTAA 1200
 DB 1141 TTGCTAGCTTTTGGCGCTTCAATCATATGGAAGAGAGAGAAATCAATACCTTAA 1200
 QY 1201 GTGGTAAATAGATGAGGGGTGCTCATCTCAATCAATTAAGGATGAGTGCATTAAT 1260
 DB 1201 GTGGTAAATAGATGAGGGGTGCTCATCTCAATCAATTAAGGATGAGTGCATTAAT 1260
 QY 1261 GAAATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 DB 1261 GAAATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 QY 1321 TTTTGGGGTAA 1332
 DB 1321 TTTTGGGGTAA 1332

RESULT 2
 AAA49202
 ID AAA49202 standard; cDNA; 1217 BP.
 XX
 XX AAA49202;
 AC
 AC 12-DEC-2000 (first entry)
 DT
 XX
 XX Soybean putative lecithin:cholesterol acyltransferase gene #1.
 DE
 XX
 XX Soybean: lecithin:cholesterol acyltransferase; phytosterol;
 KW phosphatidylcholine-sterol O-acyltransferase; heat shock; cold shock; ss.
 KM
 XX
 OS Glycine max.
 OS
 XX
 XX
 FH Key Location/Qualifiers
 FH 3..1145
 FT CDS
 FT /product="lecithin:cholesterol acyltransferase"
 FT /partial
 FT /EC_number="2.3.1.43"
 FT
 PN W0200032791-A2.
 XX
 PD 08-JUN-2000.
 XX
 PF 02-DEC-1999; 99WO-US28586.

XX 03-DEC-1998; 98US-0110782.
 PR (DUPO) DU PONT DE NEMOURS & CO E I.
 PA Cahoon RE, Kinney AJ, Sakai H, Shen JB, Butler KH, Saylor JJ;
 XX WPI: 2000-412337/35.
 DR P-PsDB; AAB01207.
 XX
 PS Claim 2; Page 38; 49pp; English.
 XX
 CC The present sequence is a putative coding sequence for a soybean
 CC lecithin:cholesterol acyltransferase (also known as
 CC phosphatidylcholine-sterol O-acyltransferase). This enzyme is found
 CC associated with high-density lipoproteins and Apolipoprotein-AI and -D.
 CC The gene and protein can be used to produce transgenic plants which have
 CC increased lipid metabolism and membrane fluidity, and therefore increased
 CC resistance to heat and/or cold shock, to alter the content of phytosterol
 CC or lecithin in grains and to identify potential herbicides.
 XX
 SQ Sequence 1217 BP; 346 A; 250 C; 262 G; 359 T; 0 other:
 Query Match 86.2%; Score 1148; DB 21; Length 1217;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 185 CTTTCATCTGGGAATCATGATGACCTCTCATCAGAGAAAAAGATGATGCTTGACACTT 244
 DB 1 CTTTCATCTGGGAATCATGATGACCTCTCATCAGAGAAAAAGATGATGCTTGACACTT 60
 QY 245 GATTGATTCAGTGTCTACTGTGCTCTCATCATGCTTTGCCAGCATGACCC 304
 DB 61 GATTGATTCAGTGTCTACTGTGCTCTCATCATGCTTTGCCAGCATGACCC 120
 QY 305 TTTCATTACCCAGCAAGACTGATGATTTCAACACTCTGGGGTTGAGACCCGGTCC 364
 DB 121 TTTCATTACCCAGCAAGACTGATGATTTCAACACTCTGGGGTTGAGACCCGGTCC 180
 QY 365 CTCACCTTGGTCCACCAACTCTCTCTCTATCTCAATCCCTGCTCAAGCATATACCG 424
 DB 181 CTCACCTTGGTCCACCAACTCTCTCTCTATCTCAATCCCTGCTCAAGCATATACCG 240
 QY 425 GATACATGGACCCCTGTAGATTCATACAAAAGCTGGGTAGCGTATGGTGGAGCTC 484
 DB 241 GATACATGGACCCCTGTAGATTCATACAAAAGCTGGGTAGCGTATGGTGGAGCTC 300
 QY 485 TGTTTGAGAGCCCTTATGACTTTAGATATGCTAGCTGTGAAGTCAACCCCTTCAAG 544
 DB 301 TGTTTGAGAGCCCTTATGACTTTAGATATGCTAGCTGTGAAGTCAACCCCTTCAAG 360
 QY 545 TGGGTTCACAAATGCTCAAAAGATCTAAAGATTTGATGAAGAACCAATTTCCAAAT 604
 DB 361 TGGGTTCACAAATGCTCAAAAGATCTAAAGATTTGATGAAGAACCAATTTCCAAAT 420
 QY 605 ATGGGAAGCCAGTGATCTTCTCTCCACAGTTTGAAGGCGCTATTTGCCACAACTAC 664
 DB 421 ATGGGAAGCCAGTGATCTTCTCTCCACAGTTTGAAGGCGCTATTTGCCACAACTAC 480
 QY 665 TAAATAGAAGACCCCTCTTGGGCAAAAAATTCATCAACACTTTCCTTTTGG 724
 DB 481 TAAATAGAAGACCCCTCTTGGGCAAAAAATTCATCAACACTTTCCTTTTGG 540
 QY 725 CTCATGGGGTGTCTATAGACGAATGTACACCTTGGCATCTGGCAACACTTTGGGAG 784
 DB 541 CTCATGGGGTGTCTATAGACGAATGTACACCTTGGCATCTGGCAACACTTTGGGAG 600
 QY 785 TGCCCTAGTGAGACCTTTATTAGTGAAGATGAACAAAGAGCTCCGAGACTTACCTT 844

DB 601 TGCCCTAGTGAGACCTTTATTAGTGAAGATGAACAAAGAGCTCCGAGAGTAACCTT 660
 QY 845 GGCCTTTGCCCTAACCCAAAAATTTTGGTCTGCAAAAACCAATAGTATCTCCAAATTA 904
 DB 661 GGCCTTTGCCCTAACCCAAAAATTTTGGTCTGCAAAAACCAATAGTATCTCCAAATTA 720
 QY 905 GGCCTTATTCAGCTCATGACATGATGTTGTTCTTAAAGACATTTGGTTTCTGAGGGG 964
 DB 721 GGCCTTATTCAGCTCATGACATGATGTTGTTCTTAAAGACATTTGGTTTCTGAGGGG 780
 QY 965 TTTATCCCTTAAAGAACACAGAAATTTACCTTGATAGGAGACATTAAGACACCAAGTGC 1024
 DB 781 TTTATCCCTTAAAGAACACAGAAATTTACCTTGATAGGAGACATTAAGACACCAAGTGC 840
 QY 1025 CTATTAATCTTATTAATGGAACCGGAGTGGACACCTTGGAAACATTTGTTATGGGAAG 1084
 DB 841 CTATTAATCTTATTAATGGAACCGGAGTGGAGACCTTGGAAACATTTGTTATGGGAAG 900
 QY 1085 GTGATTTTGAAGACGGCCAGAAATATCATATGAGGATGATGAGGAACGGTGAACCTTG 1144
 DB 901 GTGATTTTGAAGACGGCCAGAAATATCATATGAGGATGATGAGGAACGGTGAACCTTG 960
 QY 1145 TGAGCTTGTGGCGCTTCATCATCTATGGAAGAGAGAGAAAAATCAATACCTTAAAGTGG 1204
 DB 961 TGAGCTTGTGGCGCTTCATCATCTATGGAAGAGAGAGAAAAATCAATACCTTAAAGTGG 1020
 QY 1205 TTAAGATAGATGGGGTGTCTCATACTTCAATCTTAAGATGAAGTTCACCTAAATGAA 1264
 DB 1021 TTAAGATAGATGGGGTGTCTCATACTTCAATCTTAAGATGAAGTTCACCTAAATGAA 1080
 QY 1265 TAGTAGTGAGATTAATCAATTAATCTCATAGCTGAGCTGGTTAAGTAAATTTGTTT 1324
 DB 1081 TAGTAGTGAGATTAATCAATTAATCTCATAGCTGAGCTGGTTAAGTAAATTTGTTT 1140
 QY 1325 CGGGGTAA 1332
 DB 1141 CGGGGTAA 1148

RESULT 3
 AAS01081
 ID AAS01081 standard; DNA; 1299 BP.
 XX
 AC AAS01081;
 XX
 DT 31-MAY-2001 (first entry)
 DE Arabidopsis thaliana sterol acyltransferase LCAT1 DNA.
 XX
 KW Lecithin:cholesterol acyltransferase-like; LCAT; sterol acyltransferase;
 KW acyl CoA:cholesterol acyltransferase-like; ACAT; plant oil; cooking oil;
 KW nutritional supplement; dairy product; food product; salad dressing; ds.
 OS Arabidopsis thaliana.
 XX
 FH key Location/Qualifiers
 FT CDS 1..1299
 FT /*tag= a
 FT /product= "LCAT1"
 PN WO200116308-A2.
 PD 08-MAR-2001.
 XX
 PF 30-AUG-2000; 2000WO-US23863.
 XX
 PR 30-AUG-1999; 99US-0152493.
 XX
 PA (MONS) MONSANTO CO.
 XX
 PI Lassner M, Van Eenennaam A;

PR 09-MAR-1999;	99US-0123548.
PR 23-MAR-1999;	99US-0125768.
PR 25-MAR-1999;	99US-0126264.
PR 29-MAR-1999;	99US-0126785.
PR 01-APR-1999;	99US-0127462.
PR 06-APR-1999;	99US-0128234.
PR 08-APR-1999;	99US-0128714.
PR 16-APR-1999;	99US-0129845.
PR 19-APR-1999;	99US-0130077.
PR 21-APR-1999;	99US-0130449.
PR 23-APR-1999;	99US-0130510.
PR 28-APR-1999;	99US-0130891.
PR 30-APR-1999;	99US-0131449.
PR 30-APR-1999;	99US-0132048.
PR 30-APR-1999;	99US-0132407.
PR 04-MAY-1999;	99US-0132484.
PR 05-MAY-1999;	99US-0132485.
PR 06-MAY-1999;	99US-0132486.
PR 07-MAY-1999;	99US-0132863.
PR 11-MAY-1999;	99US-0134256.
PR 14-MAY-1999;	99US-0134218.
PR 14-MAY-1999;	99US-0134219.
PR 14-MAY-1999;	99US-0134221.
PR 14-MAY-1999;	99US-0134370.
PR 18-MAY-1999;	99US-0134768.
PR 19-MAY-1999;	99US-0134941.
PR 20-MAY-1999;	99US-0135124.
PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135629.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 10-JUN-1999;	99US-0138847.
PR 14-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 17-JUN-1999;	99US-0139453.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 18-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139839.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140981.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 01-JUL-1999;	99US-0142154.
PR 02-JUL-1999;	99US-0142055.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143542.
PR 14-JUL-1999;	99US-0143624.

PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144085.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144884.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145085.
PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145919.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.

PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159337.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161922.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

Query Match 33.8%; Score 449.6; DB 21; Length 1495;
Best Local Similarity 62.4%; Pred. No. 1,5e-123;
Matches 786; Conservative 0; Mismatches 444; Indels 30; Gaps 4;

OY 45 ACTCAGTAAAGTAGTGTGTGATGCTGTCATTGTCATACATGTGGGCAAGCAA 104
DB 111 ACTCATAGCGATACGCTGTGTGTGATGACATGACCTGATGTCATCAAGCTGTGGTAGCAA 170
OY 105 CCTCCAGCCCTTGTGATTTCTAATACAGAGTAAGGAGGAAACCAACTGAAAGCAAGTTGAC 164
DB 171 CCGTACGCTCTTGTGATGTGGTTCAGAGAAACGAGGTAAACAGCTTACGGCTGGA 230
OY 165 CAATCAGTCAAGCCCTTACTTTCATCTGCGAATCATGTAACCCCTCATCAAGAAAA 224
DB 231 CAGAGATATCAAGCCCAAGTAGTGTGTGTGATGACAGCTGTATATCCGATTCATAGAA 290
OY 225 GA---ATGGATGGTTCAAGCTTTGTTGATTCAGTGTCTATACCTTCTTCACTCA 281
DB 291 GAGTGTGATGATGTTAGCTATGTTGATGTCAGACGATGTTATGTTCCCTTCCAGCAG 350
OY 282 ATGCTTTGCCGAACGATACCTTCATTTACACCAAGAACGATGATTTACTTCAACAC 341
DB 351 GTGCTTCAACGATCGAATGATGTGTACTATGATGACCTGATTTGGTAAAGCTCTAGAGAAAA 410
OY 342 TCCCTGGGGTGGAGACCCGGGCTCCCTACCTTGTGTTCCACCACTGCTCTCTATCTCAA 401
DB 411 TCCCTGTGTGCAAAACCCGGGCTTCCTCATTTTCGTTTCGACCAATCACTTCTTACTCTCA 470
OY 402 TCCCTGTGTGCAAGCATATTCACCGGATACATGACACCCCTGCT---AGATTCTTCAAAA 458
DB 471 CCGTGTGTCGAGATGCGACATCTTACATGTAACATTTGGTAAAGCTCTAGAGAAAA 530
OY 459 GCTTGGCTACGCTGATGGTGAACATCTGTTTGGAGCCCTTATGACTTATGATAGTGT 518
DB 531 ATGCGGATATGTTAAGCAACCAATCTAGAGCTCCATATGATTTAGTAGGCGCT 590
OY 519 AGCTGCTGAAGGTACACCTTTCACAGTGGGTTCCAAAGTTCCTCAAGATCTAAGAAATTT 578
DB 591 GCGTCTTCGGGCAACCGCTCCCGTGTAGCTTCACAGATTCCTCAAGAGACTCAACAAATTT 650

OY 579 GATAGAAGAGCAAGCAATTTCCATATATGGAAGCCAGTATACCTTCTCCACAGTTT 638
DB 651 GGTGGAAGAAACTAGCAGCGAGAGCAAGAACCCAGATATCTCTCCATATGACCT 710
OY 639 AGGAGGCTATTTGCTTCAACCTACTAATATGAAGAACCCCTCTTGGCGCAAAAAATT 698
DB 711 AGGAGGCTTTGCTTCAACCTACTTCAACCGTATACACCCCTTATGAGCGGCAAGTA 770
OY 699 CATCAACACTTCATCTGCTTTCACCTCATGAGGCTGTGCTATAGAGAAATGATAC 758
DB 771 CATCAACACTTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 830
OY 759 CTTTGCATCTGCAACACTTTGGAGTGGCCCTTATGATGACCTTTATGATGAGGATGA 818
DB 831 ATTTGCTTGGCAACACACTCGGTGTCCCTTATGATTAACCTTTGCTGCTGCAAGGCA 890
OY 819 ACAGAGAGCTCGAGAGTATACCTTTGGCTTTGGCTTAACCCAAAATTTT---GGTCC 875
DB 891 TCAGAGAGCTCCGAGAGTATACCAATGCTTCTCATCTTACCAAAAGTGTTCACGACAG 950
OY 876 TCAAAACCAATAGTATCTACCTCAATTTAGGCTTATTCAGCTCATGATGATGATTT 935
DB 951 AACTTAACCGCTTTGCTGCTACTCCAGGTATACCTACACAGCTTACAGATGATGATGTT 1010
OY 936 TCTAAAGACATTTGTTTCTGAGGAGGTTATCTTATGAAACGCAATTTCAACCTT 995
DB 1011 TTTTGAGACATTTGATTTCTCAAGAGAGTGTGCTTCAAGAGCAAGAGTGTGCTT 1070
OY 996 GATAGGAGCATTAAGAGCAACAGTGTCTTAACTTATATGAGGAGACGGAGTGGG 1055
DB 1071 AACAGAGAGCTGATGACCTCGGAGGTGCTACCTTGCATATATGAGAGGAGTGA 1130
OY 1056 AACCTTGAACATTTGTTTATGGAAGAGTATTTGATGAGAGCGCAAGATATATCA 1115
DB 1131 TACACCGAGTTTGTATGATATGAGAAAGAGATTTGATTAAGCAACGATATTAAGTA 1190
OY 1116 TGGGATGATGAGAACGTTGACTTGTGAGCTTTGTTGGCTTCAATCAATGGA 1175
DB 1191 TGGAGATGAGATGAGGACGTTAATTTGGCAGCTTACGACGCTT----- 1235
OY 1176 AGAGAGAAATATCAATACCTTAAGTGTTAAGATAGATGGGCTGCTCATCTTCAAT 1235
DB 1236 -----GAAAGTGCATAGCTTGAACACCGTAGAGATGATGAGGTTGCAATATCTAT 1289
OY 1236 ACTTAAGATGAAGTTGCTCAATGAATTAAGTATAGTGAATTTCAATTAATTTCA 1295
DB 1290 ACTTAAGAGCATGCTGCACTTAAGAGATTTGAAGCAGATTTCAATTAATTAATGA 1349

RESULT 5
AAC39845
ID AAC39845 standard; DNA; 1498 BP.
XX
AC AAC39845;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 26113.
XX
KM Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139839.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147132.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

```

PR 06-OCT-1999: 9905-0157865.
PR 07-OCT-1999: 9905-0158029.
PR 08-OCT-1999: 9905-0158232.
PR 12-OCT-1999: 9905-0158369.
PR 13-OCT-1999: 9905-0159293.
PR 13-OCT-1999: 9905-0159294.
PR 13-OCT-1999: 9905-0159295.
PR 14-OCT-1999: 9905-0159329.
PR 14-OCT-1999: 9905-0159330.
PR 14-OCT-1999: 9905-0159331.
PR 14-OCT-1999: 9905-0159637.
PR 14-OCT-1999: 9905-0159638.
PR 18-OCT-1999: 9905-0159584.
PR 21-OCT-1999: 9905-0160741.
PR 21-OCT-1999: 9905-0160767.
PR 21-OCT-1999: 9905-0160768.
PR 21-OCT-1999: 9905-0160770.
PR 21-OCT-1999: 9905-0160814.
PR 21-OCT-1999: 9905-0160815.
PR 22-OCT-1999: 9905-0160980.
PR 22-OCT-1999: 9905-0160981.
PR 25-OCT-1999: 9905-0160989.
PR 25-OCT-1999: 9905-0161404.
PR 25-OCT-1999: 9905-0161405.
PR 25-OCT-1999: 9905-0161406.
PR 26-OCT-1999: 9905-0161359.
PR 26-OCT-1999: 9905-0161360.
PR 26-OCT-1999: 9905-0161361.
PR 28-OCT-1999: 9905-0161920.
PR 28-OCT-1999: 9905-0161992.
PR 28-OCT-1999: 9905-0161993.
PR 29-OCT-1999: 9905-0162142.

```

Query Match 33.6%: Score 448; DB 21; Length 1498;
 Best Local Similarity 62.3%: Pred. No. 4,4e-123;
 Matches 755; Conservative 0; Mismatches 445; Indels 30; Gaps 4;

```

QY 45 ACTCACAGTTACAGTAATGTTGATGCTGTCATGCTATGACATGTCAGGCGAAGCAA 104
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 111 AGCATAGCATATGCTGTTGTCGAGATGACCTGATGTCAGCTGTGGTATGCAA 110
QY 105 CCTCGACCTTTGATGTTATACAGAGTAACGAGGAGCAACCACTAGAACAGTTGAC 164
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 171 TGTGTATCCCTTTATGTTGTTCCAGGAAAGGAGGTAAACAGCTAGAGGCTGGA 230
QY 165 CAATCAGTACAGCCCTCTACTTTCATCTGCAATCATGTTACCTCTCATCAAGAAA 224
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 231 CAGAGATACAGCCAGTAGTGTGTGTGTACAGCTGTTATATCCGATTATAGAA 290
QY 225 GA---ATGATGTTACAGACTTTGTTGATTCAGTCTATGCTTCCTTCACTCA 281
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 291 GATGTGATGATGTTAGCTATGCTATGTCATGACAGCTGTATGTCCTCTTCAACG 350
QY 282 ATGCTTTGCCGAGCAGTACCTTCATTTACCAAGCAAGAACTGATGATTACTTCAAC 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 351 GTGCTTTACGAGTACGATGATGTTGATGATGACCCGATTTGATGATTCACAAATGC 410
QY 342 TCCTGGGTTGAGACCGGATCCCTCACTTTGTTCCACCAACCTCTTCTATGCA 401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 411 TCCTGGTGTCCAAACCCGGTTCCTCATTTTCGTTGACCAAAATCACTTTTATACCTCA 470
QY 402 TCCTGCTTCACAGATACCGGATACATGCGACCCCTGGT---AGATTCAATTAACAAA 458
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 471 CCCTGCTCTCGAGATGCCATCTTACATGGAATTTTGGTAAAGCTCTTAAGAAAA 530
QY 459 GCTTGGCTACGCTGATGCTGAGACTCTGTTTGAGCCCTTATGACTTATGATATGCTCT 518
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 531 ATCCGGATATGTTAAGCAACCAATCTAGAGCTCCATATGATTTCAAGTACGGCT 590
QY 519 AGCTGTGAAGTACACCTTCAACAAGTGGTGTCCAAAGTTCTTCAAGATTAAGAATTT 578
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 591 GGTGTGCTGGGCGCCCGCTCGGTAGCTCAGAGTTCTTCAAGACCTCAAAAGATTT 650

```

```

QY 579 GATAGAAGAGCAAGCAATTTCCATATGAGGAGCCAGCTGATCTTCTGCCACAGTTT 638
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 651 GTTGGAAAAAACTAGCAGGAGAGAAAGAAAGCAAGATATCTCTCCCATATGCCCT 710
QY 639 AGGAGGCTATTTGTTGCTTACACTACTAAATAGAAAACCCCTCTTGGCGCAAAATTT 698
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 711 AGGAGACTTTTGTCTCTCATTTCTCAACCCGTACACCCCTTCATGCGCGCAAGTA 770
QY 699 CATCAAACTATCTATGCTCTTCACTGCTAGGAGGTTGGTGTATGAGCAATGATAC 758
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 771 CATCAAACTATTTGTTGCTACCTGCTGCTGCTATGAGGATCTCTCATGATGAGAC 830
QY 759 CTTTGATCTGACACACTTTGGAGTGCCTTACAGTGAACCTTTATATGAGAGGATGA 818
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 831 ATTTGCTTGGCAACACACTGCTGCTGCTTATGTTAACCTTTCTGTCGACAGGCA 890
QY 819 ACAAGAGTCTCGAGAGTATACCTTTGGCTTTTGCTTAACCCAAATTTT---GTC 875
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 891 TCAGAGGACCTCCGAGAGTAACCAATGCTACTTCCATCTACCAAGTGTTCACACAG 950
QY 876 TCAAAACCAATGATGATACCTCAATTTAGCCTTATGAGCTCATGATGATGATTT 935
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 951 AACTAAACCGCTTGTCTGTAATCTCCAGGTTAATCAACAGCTTACAGATGATGCTTT 1010
QY 936 TCTAAAGACATTTGTTCTGTAAGGAGGTTATCTTATGAACAGAAATCTACCTT 995
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1011 TTTTGCAGATATGATTTCTCACAAAGAGATTTGCTTCAAGACAAAGTGTGCTTT 1070
QY 996 GATAGGAACTATAAGACACAACTGCTTATCTTTATTTATGGAACGGAGTGGG 1055
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1071 AACGAGAGAGCTGATGATGACGCTGAGTGCAGTCACTTCAATATATGAGAGGATTTA 1130
QY 1056 AACCTTGAACATTTGTTTATGGAAGAGTATTTGATGAAAGGCGCAAGAAATATCATA 1115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1131 TACACCGGAGTTTATGATATGAAAGAGAGATTCGATTAAGCAACAGATTAAGTA 1190
QY 1116 TGGGATGTTGATGGAACGCTGAACCTTGTGAGCTTTGGCGCTTCAATCACTATGAA 1175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1191 TGGAGATGAGATGGAGCGTTAATTTGGCAGCTTACGACGCTT----- 1235
QY 1176 AGAGAGAAAAATCATACCTTAAAGTGTAAAGATAGATGGGCTCTCATACTTCAAT 1235
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1236 -----GAAAGTCGATAGCTTGAACCGTAGAGATTTGATGAGTTTGCATATCATCTAT 1289
QY 1236 ACTTAAGATGAAGTTGACCTAATAAGATAGATGATGAGTTTGCATATCATCTATCA 1295
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1290 ACTTAAGACGATGTCACCTTAAGAGATTTATGAAGCATTTTCAATTTATATATGA 1349

```

```

RESULT 6
AAC64437
ID AAC64437 standard: DNA; 1562 BP.
XX
AC AAC64437;
XX
DT 12-FEB-2001 (first entry)
XX
DE Arabidopsis thaliana PDAT genomic DNA SEQ ID NO:10.
XX
KW PDAT; phospholipid; diacylglycerol acyltransferase; triacylglycerol;
XX TAG; EST; expressed sequence tag; fatty acid; oil content; ds.
XX Arabidopsis thaliana.
XX
PN W0200060095-A2.
XX
PD 12-OCT-2000.
XX
PE 28-MAR-2000; 2000MO-EP02701.
XX
PR 01-APR-1999; 99EP-0106656.
PR 10-JUN-1999; 99EP-0111321.
PR 07-FEB-2000; 2000US-0180687.

```


Best Local Similarity 61.4%: Pred. No. 2.4e-78; Matches 527; Conservative 0; Mismatches 308; Indels 24; Gaps 2;

```

OY 440 TGGTGATTCATTACAAAGCTTGCTGACGTCGATGAGACCTGCTGTTGAGCCCTT 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 694 TGAAGCTCAGAGAAAAATGCGGTATGTTACGACCAACCACTCTAGAGCTCCAT 753
OY 500 ATGACTTTAGATATGCTAGCTGCTGAAGGTCAACCTTCACAGTGGGTTCCAGTCC 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 754 ATGATTTAGATAGCGCTGCGCTGCTGGCGCCACCCGTCGCCGTAGCTACAGTCC 813
OY 560 TCAAGATCTAAAGATTGATAGAAAGCAAGCAATTCATATATGGGAAGCCAGTGA 619
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 814 TACAGACCTCAAGCAATTGTTGGAAGAAAGCTAGCAGCAAGCAAGAGAAAGCCAGTGA 873
OY 620 TACTTCTCCACAGATTGAGAGGCTTATTTGCTTACAGTCAATTAATAAGAAACCCG 679
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 874 TACTCTCTCCCATAGCCATAGGAGGACTTTCTCTCTCCATTTCCCAACCGTACACCC 933
OY 680 CCTCTTGGCGCAAAAAATTCATCAACACTTTCATTCCTTTTCACTGCAAGGGGTG 739
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 934 CTTCATGCGCGCCGCAAGTACATCAAACTTTGTTGCTACCTCCGCGCCATGGGGTGA 993
OY 740 CTATAGACAAATGATACACTTTGCTGTCGCAACACTTTGGGAGTCCCTAGTGGACC 799
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 994 CGATCTCTCAGATGAGACATTTGCTTGGCAACACACTCGGTGCTCCCTTAACTAAC 1053
OY 800 CTTTATTAAGTAGAGATGAACAAAGAGCTCCGAGATGACCTTGGCTTTGGCTTAAC 859
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1054 CTTTGCTGTCTACAGCCGCACTCAGAGAGCTCCGAGATGACCAATGCTACTTCCATCTA 1113
OY 860 CAAAAATTTTGGCTCTCA---AAACCAATAGTATTAACCTCCAAATAGCCCTTATTCAG 916
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1114 CCAAAGTGTTCACGACAGCAACTAAACCGCTTGTCTGCTACTCCCGCAGTAACTACACA 1173
OY 917 CTCATGACATGTTGATTTTCTTAAAGACATTTGCTTCTCTGAGGGTTTATCTTATG 976
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1174 CTTCACAGATGATCGGTTTTTTCAGACATTTGATTTCTCACAAAGAGTTTGCTTTACA 1233
OY 977 AAACAGATTTCTACCTTGATAGGAAACATTAAGCACCACAGACCTTAACTATCTGA 1036
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1234 ACACAGAGTGTTCCTTTTACACAGAGAGCTGATGATCCGGGAGTCCAGTCACTTGA 1293
OY 1037 TTTATGGGAAGCGGAGTGGGAACCTTGGAAACATTTGTTATGGGAAGTGTATTTGATG 1096
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1294 TATATGGGAGAGAGTGTGATACCGGAGTTTGTATGATGAAAGAGAGATTCGATGA 1353
OY 1097 AACGCCCAAAATATCATATGATGGGATGCTGATGGAACGCTGAACCTTGGTACTTGTGG 1156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1354 ACCAACCCAGATTTAAGTATGGAGATGGAGTGGCGGTTAATTTGGGACCTTAGCAG 1413
OY 1157 CGCTTCATCATATGAGAAAGAGAAATCAATACCTTAAAGGTGTTAAGATGATG 1216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1414 CTTT-----GAAAGTGCATAGCTTTGAACACCCGATAGAGATTGATG 1452
OY 1217 GGGTGTCTCATCTCAATACTTAAGAGATGAAGTTGCACATAAATAGTAGGTGAGA 1276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1453 GAGTTTCCGATACATCTATTAATAAGAGAGATGCACTTAAAGAGATTATGAGCAGA 1512
OY 1277 TTTACTCAATTAATTTCTCA 1295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1513 TTTCAATTAATTAATGCA 1531

```

RESULT 8

AA501087
ID AA501087 standard; cDNA; 264 BP.
AC AA501087;
XX
DT 31-MAY-2001 (first entry)
XX
DE Soybean sterol acyltransferase LCAT EST sequence #1.

```

XX Lecithin:cholesterol acyltransferase-like; LCAT; sterol acyltransferase;
KW acyl CoA:cholesterol acyltransferase-like; ACAT; plant oil; cooking oil;
KW nutritional supplement; dairy product; food product; salad dressing;
KW soybean; Arabidopsis thaliana; expressed sequence tag; EST; ss.
XX
OS Glycine max.
XX
XX WO200116308-A2.
XX
XX 30-AUG-2000; 2000MO-US23863.
XX
XX 30-AUG-1999; 99US-0152493.
XX
XX (MONS ) MONSANTO CO.
XX
XX Lassner M, Van Benennaam A;
XX
XX WPI; 2001-169010/17.
XX
XX
XX New isolated nucleic acid encoding plant lecithin:cholesterol
PT acyltransferase-like or acyl (coenzyme A) CoA:cholesterol
PT acyltransferase-like polypeptides, for modifying the sterol content and
PT oil production of plants -
XX
XX
XX Claim 5; Page 86; 127pp; English.
XX
XX The present sequence for soybean LCAT (lecithin:cholesterol
CC acyltransferase-like) EST sequence #1 is closely related to the
CC Arabidopsis thaliana LCAT1 sequence. Several novel polynucleotides
CC encoding the plant sterol acyltransferases LCAT (AA501081-AA501104,
CC AA501341) and ACAT (acyl CoA:cholesterol acyltransferase-like;
CC AA501311-AA501319) are described. A yeast LCAT related open reading
CC frame, LK01 gene sequence (AA501342), and a rat ACAT (AA501105) cDNA
CC sequence are also described. The polynucleotides encoding LCAT or ACAT
CC are used to produce LCAT or ACAT polypeptides. They can also be used in
CC a recombinant construct to transform a host cell (preferably of a plant)
CC or a plant. The recombinant construct is used to increase or decrease
CC the sterol content of the host cell or plant. It can be used to alter
CC oil production of the cell or plant, preferably by increasing it. The
CC oil of the plant or the plant itself is used as a food product, or as
CC nutritional or dietary supplements, or in pharmaceutical compositions
CC for lowering cholesterol. The oil can be used in foods e.g. margarine,
CC butter, cooking oil, and dressings e.g. salad dressings, mayonnaise,
CC cheese, processed meat, pasta, sauces, cereals, desserts, dips, chips,
CC baked goods, pastries, cookies, snack bars, confections, chocolates, and
CC beverages. The alteration in sterol content and/or composition can also
CC provide a plant with tolerance to stress and insect damage.
XX
XX
XX Sequence 264 BP; 62 A; 73 C; 52 G; 74 T; 3 other;
XX
XX
XX Query Match 18.7%; Score 249; DB 22; Length 264;
XX Best Local Similarity 98.5%; Pred. No. 5e-64;
XX Matches 260; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
XX
OY 315 CCAGAACATCGATGATTAATTCACACATCTGGGGTTGAGACCGGGTCCCTACCTTGG 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 CCAAGACATCGATGATTAATTCACACATCTGGGGTTGAGACCGGGTCCCTACCTTGG 60
OY 375 TTTCACCAACTCTCTTCT-ATCTCAATCTCTGCTCAAGCATATCAACCGATACATG 433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 TTTCACCAACTCTCTTCTCTCATCTCATTCATCCGCTCAAGCATATCAACCGATACATG 120
OY 434 CACCCCTGTAGATTCATTACAAAGCTTGGCTACCGCTGATGTTAGAGCTCTGTTGGAG 493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 121 CACCCCTGTAGATTCATTACAAAGCTTGGCTACCGCTGATGTTAGAGCTCTGTTGGAG 180
OY 494 CCCCTTATGACTTTAGATATGTTGCTAGCTGCTGAAGGTACCCCTCACAAAGTGGTCCA 553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 181 CCCCTTATGACTTTAGATATGTTGCTAGCTGCTGAAGGTACCCCTTCAACAAAGTGGTCCA 240

```

QY 554 AGTCTCTCAAGATCTAAGAATT 577
 |||||
 Db 241 NGTTCCTCAAGATCTAAGAATT 264

RESULT 9
 AAA49204
 ID AAA49204 standard; cDNA: 1500 BP.
 XX
 AC AAA49204;
 XX
 DT 12-DEC-2000 (first entry)
 XX
 DE Corn putative lecithin:cholesterol acyltransferase gene #4.
 XX
 KW Corn; lecithin:cholesterol acyltransferase; phytosterol;
 XX phosphatidylcholine-sterol O-acyltransferase; heat shock; cold shock; ss.
 OS
 XX Zea mays.
 FH Location/Qualifiers
 FT 24..1277
 FT CDS /*tag= a
 FT /product= "lecithin:cholesterol acyltransferase"
 FT /EC_number= "2.3.1.43"
 XX
 PN MO200032791-A2.
 XX
 XX 08-JUN-2000.
 PD
 XX 02-DEC-1999; 99WO-US28586.
 PF
 XX 03-DEC-1998; 98US-0110782.
 PR
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 PA Cahoon RE, Kinney AJ, Sakai H, Shen JB, Butler KH, Saylor JJ;
 XX WP1: 2000-412337/35.
 DR P-PSDB: AAB01209.
 XX
 XX Polynucleotide encoding plant lecithin:cholesterol acyltransferase
 PT enzyme useful for producing transgenic plants and for producing
 PT antibodies specific to which is useful for screening cDNA expression
 PT libraries -
 XX
 PS Claim 2; Page 41-42; 49pp; English.
 XX
 CC The present sequence is a putative coding sequence for a corn
 CC lecithin:cholesterol acyltransferase (also known as
 CC phosphatidylcholine-sterol O-acyltransferase). This enzyme is found
 CC associated with high-density lipoproteins and Apolipoprotein-AI and -D.
 CC The gene and protein can be used to produce transgenic plants which have
 CC increased lipid metabolism and membrane fluidity, and therefore increased
 CC resistance to heat and/or cold shock, to alter the content of phytosterol
 CC or lecithin in grains and to identify potential herbicides.
 XX
 SQ Sequence 1500 BP: 304 A; 469 C; 454 G; 271 T; 2 other;

Query Match 15.6%; Score 208; DB 21; Length 1500;
 Best Local Similarity 51.9%; Pred. No. 2e-51;
 Matches 516; Conservative 1; Mismatches 472; Indels 6; Gaps 2;

QY 283 TGCCTTGGCGAAGCAGTACCTTCATTACACCAAGACTCGATGATTACTTCAACACT 342
 |||||
 Db 273 TGYTTCCGCGAGAGATGCGACGCGGCGGCTCCGACGACTACCGCAAGCC 332

QY 343 CCTGGGGTGGAGACCGGGTCCCTCATTGTTGCCAACCACTCTCTCTCATCTCAAT 402
 |||||
 Db 333 CAGGCGCGGAGACCCGGGCTCTTCTTGATGACACCGGGCTTCGCGTACCCGAGC 392

QY 403 CCTGCTCAAGCATATACCGGATACGACCCCTGTAGATTCAATCAAAAGCTT 462
 |||||

Db 393 CCAGACCGGAGAAACTTCTC---GTACATGACAAAGTTCGTGCGGGCTGGAGGCGTC 449
 QY 463 GGCTACGCTGATGTGAACACTGTGTTGGAGCCCTTATGACTTATGATATGTTACT 522
 |||||
 Db 450 GCGTACCGCGAGCGGAGAACCTGTTCCGCGCGCCCTACGACTTCGGGTACGCCGTCGCC 509

QY 523 GCTGAAGGTCAACCCCTTCAAGAGTTCGCAAGTTCCTCAAAAGATTAAGAAATTGATA 582
 |||||
 Db 510 CCGCCAGGCGACCGCTCGAGGCTCGCNGACCGCTTCTTGCGGGCGCTCAGAGGCTGGTA 569

QY 583 GAAAGAGCAAGCAATTCATATATGGAAGCCAGTATCTCTCCACAGTTTAGA 642
 |||||
 Db 570 GAGAGAGGAGAGCGGCGTAAAGGAGAGGCGGTTGACCATCTGCGGCGACAGCTACGCG 629

QY 643 GGCTATTGTTCTCAACACTAATATGAAGAACCCCTCTTGCGCGCAAAAATTATC 702
 |||||
 Db 630 GGCACGGTGGCGCACACTTCTACTGGGCGGCGCTTGCCGTGGCGAGCGCTTCGTC 689

QY 703 AAGACTCATGCTCTTCAAGCTCCATGGGGGTGGTGCATATAGAGAAATGTACACTT 762
 |||||
 Db 690 CGGCGGTGTCGCCGTTGCGGCGCTGGGAGGCGTCTCTTGCAATGCTGCAATC 749

QY 763 GCATCTGGCAACACTTTTGGAGTGGCCCTAGTGAACCCCTTATATGATGAGGATGAACA 822
 |||||
 Db 750 GTGCGCGGCAACATCTCGGCGCTGCGTTCGACACCGCTGGCGCTCAAGGCGAGTAC 809

QY 823 AGAAGCTCCGAGATTAACCTTTGGCTTTGCTTAACCCAAAATTGTTGCTTAATAA 882
 |||||
 Db 810 CGGAGCTGACAGAGAGAGCTTGGCGGCTGCCAACCCCAACGCAATTAGAGCCGCGGAG 869

QY 883 CCAATAGATATACATCCATTAATGACCTTATGACGTCATGATGATGATTTCTAATAA 942
 |||||
 Db 870 CCATGTTGACACACGAGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 929

QY 943 GACATTGTTTCTTGAAGGGGTTATCTTATGAAGACGAATTTCTACCTTGTATAGG 1002
 |||||
 Db 930 GGCATCGGGCTAGGCGGCAATTTGCGGTACGATCCCGGTGCTGCGCCCTGTTCCGG 989

QY 1003 AACATTAAGACACCAAGTGCCTTAACTGTATATGGAAGCGGAGTGGGAACCTTG 1062
 |||||
 Db 990 GAGCTGCCATCTCCCGGGGTGCCGTGCTGTGTCGCGGGGTGGCTGCACAGCGCG 1049

QY 1063 GAAACATTTGTTTATGGGAAAGGTATTTGATGAAGCGCCGAATATCATATGGGGAT 1122
 |||||
 Db 1050 GAGATGCTGGCCTTACCCGAGACGACTTGCAGCGAGCGCATATGCTATGGGAGAC 1109

QY 1123 GGTGATGAGAACGCTGATGAGTGTGAGCTTGTGGCGCTTCAATCACTATGGAAGAGNG 1182
 |||||
 Db 1110 GCGGACGGGCTGTACACCTGCTAGGCTCTGTCGTGACCCGTGCGGAGGCTTCCT 1169

QY 1183 AAAAATCAATACCTTAAAGTGTTAAGATAGATGGGGTGTCTCATCTTCAATTAAG 1242
 |||||
 Db 1170 ACAGCT---TACTTTAGATGCTCAAGTGCCCAACGCTGTACACAGCGGCGCTTTCGTG 1226

QY 1243 GATGAAGTTCGACTAATGAATAGTAGTGGAT 1277
 |||||
 Db 1227 GAGGATGCTGCTCGCGTTATATAGCGGCAT 1261

RESULT 10
 AAS01088
 ID AAS01088 standard; cDNA: 273 BP.
 XX
 AC AAS01088;
 XX
 DT 31-MAY-2001 (first entry)
 XX
 DE Soybean sterol acyltransferase LCAT EST sequence #2.
 XX
 KW Lecithin:cholesterol acyltransferase-like; LCAT; sterol acyltransferase;
 KW acyl CoA:cholesterol acyltransferase-like; ACAT; plant oil; cooking oil;
 KW nutritional supplement; dairy product; food product; salad dressing;
 KW soybean; Arabidopsis thaliana; expressed sequence tag; EST; ss.

```
XX Glycine max.
OS
XX WO200116308-A2.
XX
XX 08-MAR-2001.
XX
XX 30-AUG-2000: 2000WO-US23863.
XX
XX 30-AUG-1999: 99US-0152493.
XX
XX (MONS ) MONSANTO CO.
XX
XX Laasner M, Van Benennaam A:
XX
XX WPI: 2001-169010/17.
XX
XX
XX New isolated nucleic acid encoding plant lecithin:cholesterol
PT acyltransferase-like or acyl (coenzyme A) CoA:cholesterol
PT acyltransferase-like polypeptides, for modifying the sterol content and
PT oil production of plants -
XX
XX Claim 5; Page 87; 127pp; English.
XX
XX The present sequence for soybean LCAT (lecithin:cholesterol
CC acyltransferase-like) EST sequence #2 is closely related to the
CC Arabidopsis thaliana LCAT1 sequence. Several novel polynucleotides
CC encoding the plant sterol acyltransferases LCAT (AA5001081-AA501104,
CC AA501341) and ACAT (acyl CoA:cholesterol acyltransferase-like;
CC AA501311-AA501319) are described. A yeast LCAT related open reading
CC frame, LK01 gene sequence (AA501342), and a rat ACAT (AA501105) cDNA
CC are also described. The polynucleotides encoding LCAT or ACAT
CC are used to produce LCAT or ACAT polypeptides. They can also be used in
CC a recombinant construct to transform a host cell (preferably of a plant)
CC or a plant. The recombinant construct is used to increase or decrease
CC the sterol content of the host cell or plant. It can be used to alter
CC oil production of the cell or plant, preferably by increasing it. The
CC oil of the plant or the plant itself is used as a food product, or as
CC nutritional or dietary supplements, or in pharmaceutical compositions
CC for lowering cholesterol. The oil can be used in foods e.g. margarine,
CC butter, cooking oil, and dressings e.g. salad dressings, mayonnaise,
CC cheese, processed meat, pasta, sauces, cereals, desserts, dips, chips,
CC baked goods, pastries, cookies, snack bars, confections, chocolates, and
CC beverages. The alteration in sterol content and/or composition can also
CC provide a plant with tolerance to stress and insect damage.
CC
XX Sequence 273 BP; 80 A; 56 C; 62 G; 69 T; 6 other:
SQ
Query Match 15.5%; Score 206; DB 22; Length 273;
Best Local Similarity 97.7%; Pred. No. 3.5e-51;
Matches 217; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
OY 1 ATGAGAGAGCAACAGAGAGGCTTCAGATTGAGTGTGACATCTACAGTTACAGTA 60
DB 53 ATGAGAGAGCAACAGAGAGGCTTCAGATTGAGTGTGACATCTACAGTTACAGTA 112
OY 61 GTTGTGTGATCTGTCTATTGCTATGACATGTGGGCAAGCAACCTCGACCTTTGATT 120
DB 113 GTTGTGTGATCTGTCTATTGCTATGACATGTGGGCAAGCAACCTCGACCTTTGATT 172
OY 121 CTAATACACGAGTACGAGGAGCAACCACTAGAGCAAGGTGACCATAGTACAGCC 180
DB 173 CTAATACACGAGTACGAGGAGCAACCACTAGAGCAAGGTGACCATAGTACAGCC 232
OY 181 TCTACTTTCATCTGGCAATCATGATGATCCTCTCATCAAGAAA 222
DB 233 TCTACTTTCATCTGGC-ATCTGTGATCCTCTCANNAAGANA 273
```

```
RESULT 11
AAA69662
ID AAA69662 standard: cDNA; 563 BP.
XX
```

```
AC AAA69662:
XX
XX 08-NOV-2000 (first entry)
XX
XX Pinus radiata lecithin:cholesterol acyl transferase cDNA SEQ ID NO:136.
DE
XX
XX Eucalyptus grandis; Pinus radiata; modification: isoprenoid; plant;
XX metabolism; isoprenoid biosynthetic pathway; terpenoid; steroid;
XX genome mapping; physical mapping; positional cloning; forestry;
XX agriculture; medicine; fermentation; plant development; pest resistance;
XX pinene; myrcene; Monterey pine; ss.
XX
XX Pinus radiata.
XX
XX WO200036081-A2.
XX
XX 22-JUN-2000.
XX
XX 16-DEC-1999: 99WO-NZ00219.
XX
XX 17-DEC-1998: 98US-0215504.
XX
XX 29-JUL-1999: 99US-0146441.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
XX Havukkala IJ;
XX
XX WPI: 2000-431575/37.
XX P-PSDB: AAB18062.
XX
XX New plant polynucleotides encoding polypeptides involved in the
PT production and modification of isoprenoids, useful in forestry and
PT agriculture for manipulation of isoprenoid metabolism -
XX
XX Claim 1; Page 94; 164pp; English.
XX
XX The present invention describes plant polynucleotides encoding
CC polypeptides involved in the production and modification of isoprenoids,
CC such as terpenoid and steroid compounds. The polynucleotides are used
CC in genome mapping, in physical mapping and in positional cloning of
CC genes. The polynucleotides and polypeptides are useful in forestry and
CC agriculture for manipulation of isoprenoid metabolism, in medicine for
CC therapeutic effects, including direct application in diseased organisms
CC or indirect application by transgenic organisms and in fermentation and
CC chemical processing industries involving isoprenoids. In plant
CC applications, manipulating isoprenoid pathways or isoprenoid composition
CC may, for example, affect plant development, pest resistance, and the
CC value of extractives (e.g. pinene and myrcene). The ubiquitous and
CC varied roles of isoprenoids make the polynucleotides attractive targets
CC for biotechnical applications in a variety of fields. AAA69527 to
CC AAA69690 and AAB18004 to AAB18143 represent Eucalyptus grandis and Pinus
CC radiata polynucleotides and proteins used in the exemplification of the
CC present invention.
XX
XX Sequence 563 BP; 163 A; 114 C; 137 G; 149 T; 0 other:
SQ
Query Match 10.8%; Score 144; DB 21; Length 563;
Best Local Similarity 58.7%; Pred. No. 1.6e-32;
Matches 296; Conservative 0; Mismatches 190; Indels 18; Gaps 2;
OY 87 CACATGTGGGGAGCAACACCTCGACCTTTGATTCTAATACAGGTAACGAGGAACCA 146
DB 76 CATATTTCACAGCAGTAATATCCACCTGTAGTAGTCACGACACGAGGAATCA 135
OY 147 ACTAGAGCAAGGTGACCAATCATGACAGCCCTCTACTTTCATCTGCGAATCATGTA 206
DB 136 GGTTCAGCAAGGCTTAACCTGACAGACTATTAACCCAGTGGCTGTTCGAGAGGCG-- 193
OY 207 CCTCTCATCAAGAAAAAGATGATGATGTTGATGTTGATTCAGAGTGTCTACT 266
DB 194 -----AATTGGGAGAGGAGTGGTTTCAGAAATATGTTGATGTTCTGTCTCT 243
```


XX Cahnou RE, Kinney AJ, Sakai H, Shen JB, Butler KH, Saylor JJ;
PI WPI: 2000-412337/35.
XX P-PSDB: AAB01205.
XX
PT Polynucleotide encoding plant lecithin:cholesterol acyltransferase
PT enzyme useful for producing transgenic plants and for producing
PT antibodies specific to which is useful for screening cDNA expression
PT libraries
XX
PS Claim 2: Page 35-36; 49pp; English.
XX
CC The present sequence is a putative coding sequence for a corn
CC lecithin:cholesterol acyltransferase (also known as
CC phosphatidylcholine-sterol O-acyltransferase). This enzyme is found
CC associated with high-density lipoproteins and Apolipoprotein-A1 and -D.
CC The gene and protein can be used to produce transgenic plants which have
CC increased lipid metabolism and membrane fluidity, and therefore increased
CC resistance to heat and/or cold shock, to alter the content of phytoosterol
CC or lecithin in grains and to identify potential herbicides.
XX
SQ Sequence 542 BP; 84 A; 179 C; 181 G; 90 T; 8 other;

Query Match 7.9%; Score 105; DB 21; Length 542;
Best Local Similarity 53.5%; Pred. No. 6,9e-21;
Matches 219; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

OY 629 CCACAGTTAGAGGCGCTATTTGCTTACACTACTAATAAAGAACCCCTCTTGGC 688
DB 5 CGCACAGCTACGCGCGACGCTGGCGCACAGTCTCTACTGCGCGCGCTTGCCTGGC 64
OY 689 GCAAAAATTCATCAAAACATTCATTTCTTCAAGCTCCATGGGGTGTATAGAG 748
DB 65 GCAGCGGCTTCGTCGCGGCTTCGTCGCGCTGCGCGACGCGGAGGCGTCCTTG 124
OY 749 AATGTACACCTTTGATCTGGACACTTTGGAGTCCCTAGTGCACCTTTATAG 808
DB 125 GCATGCTGCAATCTGTCGCGGCAACATCTCGGCTTCTGCTGACCGCTGGCG 184
OY 809 TGAGGATGAACAAAGAACCTCGAGAGTAACTTGGCTTGGCTTACCAACCAATTT 868
DB 185 TCAAGGGCAGTACCGAGCCTGCGACAGCCTCTGGCGCTGCCAACCACCGCAT 244
OY 869 TTGGTCTCTAAAACCAATATGATATCAATTAAGGCTTATTCAGCTCATGACATG 928
DB 245 TTAGAGCCGCGAGCAGTGTGACACAGGAGGAGGAGTACAGCGCCACGACATG 304
OY 929 TTGATTTTCTAAAAGCATTTGTTTCTGAAAGGGTTTATCTTATGAACACGATTG 988
DB 305 CGGACTTCTCGACGCGCATCGGCGTAAAGGCGCAATTTGCGCGTACGATCCGCGTGC 364
OY 989 TACCTTGTAGGAACATTAAGACACGACAGTGGCTTAATCTTAT 1037
DB 365 TCCCTCTGTTCCGGAGCTGCATCTCCGCGGTGCGCTGCTTGT 413

RESULT 14
AAAA9205
ID AAA49205 standard: cDNA; 1660 BP.
XX
AC AAA49205;
XX
DT 12-DEC-2000 (first entry)
XX
DE Corn putative lecithin:cholesterol acyltransferase gene #5.
XX
KM Corn; lecithin:cholesterol acyltransferase; phytoosterol;
XX phosphatidylcholine-sterol O-acyltransferase; heat shock; cold shock; ss.
OS
XX Zea mays.
FH Key Location/Qualifiers

FT CDS 77..1396
FT /*tag= a
FT /product= "lecithin:cholesterol acyltransferase"
FT /EC_number= "2.3.1.43"
XX
XX WO200032791-A2.
XX
XX 08-JUN-2000.
XX
XX 02-DEC-1999; 99WO-US28586.
XX
XX 03-DEC-1998; 98US-0110782.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Cahnou RE, Kinney AJ, Sakai H, Shen JB, Butler KH, Saylor JJ;
PI WPI: 2000-412337/35.
XX P-PSDB: AAB01210.
XX
XX
XX Polynucleotide encoding plant lecithin:cholesterol acyltransferase
XX enzyme useful for producing transgenic plants and for producing
XX antibodies specific to which is useful for screening cDNA expression
XX libraries
XX
PS Claim 2: Page 43-44; 49pp; English.
XX
CC The present sequence is a putative coding sequence for a corn
CC lecithin:cholesterol acyltransferase (also known as
CC phosphatidylcholine-sterol O-acyltransferase). This enzyme is found
CC associated with high-density lipoproteins and Apolipoprotein-A1 and -D.
CC The gene and protein can be used to produce transgenic plants which have
CC increased lipid metabolism and membrane fluidity, and therefore increased
CC resistance to heat and/or cold shock, to alter the content of phytoosterol
CC or lecithin in grains and to identify potential herbicides.
XX
SQ Sequence 1660 BP; 361 A; 475 C; 499 G; 325 T; 0 other;

Query Match 7.2%; Score 95.4; DB 21; Length 1660;
Best Local Similarity 46.9%; Pred. No. 8,8e-18;
Matches 574; Conservative 0; Mismatches 606; Indels 45; Gaps 7;

OY 95 GGGCAAGCAACCTCGACCTTTGATTTAATACAGTACGAGGAGGAACCACTAGAG 154
DB 177 GCGCCAACTCTCCACCCATCTTCTGTAGTGGGTGAGAGCGACCTGGAG 236
OY 155 CAAGGTGACCAATAGTCAAGCCCTTACTTATCTGCGAATCATGTACCTCTCA 214
DB 237 CACGCTTACCGAGAGTACCGGCGCTGCGCCGACTGC-----GCG 281
OY 215 TCAGAAAAAGAAATGATGTTCAAGCTTTGGTTGATTCAGTGTCAATCTTGTCT 273
DB 282 CGATGAAGGGGAAGGGGTGTCTGTGTGAAGAAGTTGCGAGCTGTCTCGT 341
OY 274 --TTCACTCAATGCTTTGGCGAAGCATGACCTTCAATTAACCAAGAATCGATG 331
DB 342 ACTAGTGCAGTCTTGTGAGAGAGATGAGCTGTCTGACAGCCCTGATCAAGAGT 401
OY 332 ACTTCAACACTGCTGGGTGTGAGACCGGCTCCATCTTGTGTGCCAAGCTCTTC 391
DB 402 ACCGGAACCTCGCGCGGTGAGAGCGGATGCCAATCTGCGGTCCAAAGACCTTCA 461
OY 392 TCTATCTCAATCTGCTCAAGCATATACCGGATATGACGACCCCTGATGATTGAT 451
DB 462 GCCAAGAAGACCCCTCAAGTCAG-----CTGTTGCTCTGGAAGAGTGAAGCGGAC 515
OY 452 TACAAAAGCTTGGCTACGCTGATGTGAGACTCTTTGGAGCCCTTATGACTTTAGAT 511
DB 516 TGAAGACATGGGATACCGAGAGGAGACACATGTTGGAACCCCTTACGACTTCGCT 575
OY 512 ATGCTCTAGCTGCTGAAGTCACTTCAAGATGGGTTCCAAAGTTCTCAAAAGATCTAA 571
DB 576 ACGCGCGCGCTGCCCGCGCGAGCTCGAGGTGTACTCCGCTACTTCAAGAGAGCTGA 635

QY	572	AGAAATTTGATAGAAAGAACCAATTCCAAATATGGGAAGCCAGATACCTTCTGCC	631
Db	636	TGGAGCTGGTCGAGGCCGAGGACGAAGAGACCCGGAAGAGCCGTATCTCGGCCACA	695
QY	632	ACAGTTTAGAGGCCCTAATTTGTCCTCAACTACTTAATATGAAACCCCCCTCTTGCGCA	691
Db	696	GGTTGGGGCGCATGTCGTGGGCTCGAATTCGTCCGGAACAATCCGGCCGGCGGGGGCGG	755
QY	692	AAAAATTCATCAAAACACTTCATTCGTCCTTTCACTCCATGGGGTGTGCTATAGAGAA	751
Db	756	ACACACATCGAGGCGCTCGTCCTGTGGCGCCGACGCTCCCGGGGGCTTCTGGAACGG	815
QY	752	TGTACACCTTTCGATCTGCA---ACACTTTGGGAGTGGCCCTAGTGGACCTTTATTAG	808
Db	816	TCCGCACTTCGCGTCCGGACGCAATCTTACGTCGACGACGACGCGCGTGCGCA	875
QY	809	TGAGGGATGAAACAAAGAACCTCCGAGATTAACCTTTGGCTTTTGCTTAACCCAAATTT	868
Db	876	CGCGAGCCATGTGAGAGAGCTTCGAGAGCCCAATCGGAACCTTCCCTCGCCGCGGT	935
QY	869	TTCGTCCTCAAAAACCAATAGTATTAATCTCAATTAAGCCTTA-----TTCAGCTC	919
Db	936	TGCGGCGCTGCAAGGCGCGCTCGTGGTCAACAGGAGCGGAATCTACCTCCGCGTCCG	995
QY	920	ATGACATGCTGATTTTCTTAAAGACATTTGTTTCTCTGAAGGGGTTATCTTATGAA	979
Db	996	ACGACATGAGGCGCTTCCTCGCCGCTGCGCTCCGCGAGGCGCGAGCCCTTCAGGA	1055
QY	980	CACGAATTTACCTCTGATAGGAGACATTAACACACACAAGTGCCTTAATCTGTATTA	1039
Db	1056	GACGGGCGCTCCCAAGATGGCAGCTTCGCGCGCCGATGTGGCCCATGACGTACATCA	1115
QY	1040	TGGGAACGGGAGTGGGAACCTTGGAACATTTGTTTATGGGAAAGGTGATTTTGATGAAC	1099
Db	1116	GGGGGTGGCAACAGAGCGCGCTCGGCTGGTTCGTGGGGGAAGACTTCGACCGCG	1175
QY	1100	GGCCAGAAATATCA---TATGGGATGTGATGGAACGATGAACTTGGTGAAGCTTGTGG	1156
Db	1176	CCCGGAGGTGGCGGCGTACGGGGACGAGATGGCAAGATCAATTTGATCAGCGCTTGG	1235
QY	1157	GGCTT-----CAATCAGTATGGAAGAGGAGGAAAAATCAATACCTTAAGCTGTTAAGA	1210
Db	1236	CGTTTGAGAGGAGATGCTCGGCGACCCGAGGACAGAAAGACAGTTCAATCCATCAGA	1295
QY	1211	TGATGGGGTGTCTCATCTTCAATCTTAAATGATGAAGTGAAGTGCATTAATGAATAGTAG	1270
Db	1296	TCAATTAAGGCCAGCATTTCTAGCATGCTACGGAATGATTTTGGCCCTCAGAGGTCATTC	1355
QY	1271	GTGAGATTACTTCATTAATTTCTCA	1295
Db	1356	AAGAAATTTGTGAGCCCAATTAATCA	1380
RESULT 15			
ID	AAAA9201		
XX	AAAA9201 standard; cDNA: 921 BP.		
AC	AAAA9201;		
XX	12-DEC-2000 (first entry)		
DT			
XX	Corn putative lecithin:cholesterol acyltransferase gene #2.		
DE			
XX	Corn; lecithin:cholesterol acyltransferase; phytosterol;		
KW	phosphatidylcholine-sterol O-acyltransferase; heat shock; cold shock; ss		
XX			
OS	Zea mays.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	70..768	
FT	/tag= "a" lecithin:cholesterol acyltransferase"		
FT			

FT	/partial	
XX	/EC_number="2.3.1.43"	
XX		
PN	WO200032791-A2.	
XX		
XX	08-JUN-2000.	
PD		
XX		
PF	02-DEC-1999; 99WO-US28586.	
XX		
PR	03-DEC-1998; 98US-0110782.	
XX		
PA	(DUPO) DU PONT DE NEMOURS & CO E. I.	
XX		
PI	Cahoon RE, kinney AJ, Sakai H, Shen JB, Butler KH, Saylor JJ;	
XX		
DR	WPI: 2000-412337/35.	
XX	P-PSDB: AAB01206.	
PT		
XX		
XX		
XX		
PS		
XX		
CC	Claim 2; Page 36-37; 49pp; English.	
CC		
CC	The present sequence is a putative coding sequence for a corn	
CC	lecithin:cholesterol acyltransferase (also known as	
CC	phosphatidylcholine-sterol O-acyltransferase). This enzyme is found	
CC	associated with high-density lipoproteins and Apolipoprotein AI and	
CC	The gene and protein can be used to produce transgenic plants which h	
CC	increased lipid metabolism and membrane fluidity, and therefore increa	
CC	resistance to heat and/or cold shock, to alter the content of phytos	
CC	or lecithin in grains and to identify potential herbicides.	
XX		
XX		
SQ	Sequence 921 BP; 176 A; 300 C; 280 G; 164 T; 1 other;	
	Query Match 6.5%; Score 86.4; DB 21; Length 921;	
	Best Local Similarity 51.3%; Pred. No.3.2e-15;	
	Matches 319; Conservative 0; Mismatches 276; Indels 27; Gaps	
QY	95 GGGCAAGCAACCTGCAGACCTTGATTTCMAATACCGATACGGAGGAGGACCACTAGAG 15	
DB	170 GCGCCAAACACTTCCACCCACTTTCGTGTGATGCTGGGCTAGCTGCAGCAACCTGGAGG 22	
QY	155 CAGGTTGACCAATCAGTACACAGCCCTCTACTTTCATTTGGGAATCAATGATACCTCTCA 21	
DB	230 CACGGCTCACCAGGAGATACCGCGCGTGGTCCGCAATGG-----GCG 27	
QY	215 TCAGAAAAAGAAATGATGGTTTCAGACCTTTGGTTGATTCAGTGTCTACTTGTCTCT- 27	
DB	275 CCATGAAGGGAGGAGGGGTTCGTGGTGTGTGGAAGACATTCGGAGCTGCTGTCTGTG 33	
QY	274 --TTCACTCAATGCTTTGCGCAAGCATGACCTTCATTTACCAACCAACTGCATGATT 33	
DB	335 ACTAGCTGAGAGCTTCAGGAGAGCAAGATGAGCTCTGTACGACCTCTCCATCAACGAGT 39	
QY	332 ACTTCACACACTCTCTGGGGTTGAGACCGGGTCCCTCACTTTGGTTGCTCCACCAACTCTTC 39	
DB	395 ACCGGAACCTCGCGCGCTCGAGACGCAAGTGTCCCAACTTGGCTCCCAAGAGCTTTCA 45	
QY	392 TCTATCTCAATCTCTGTCTCAAGCATATCACCGATACATGGACCCCTGGTAGATTTCAT 45	
DB	455 GCCACAGAACCCCTCAAGTCAGA-----CTGGTGCCTCGGAAGAGCTGAGAGCGGAC 50	
QY	452 TCAAAAGCTTGGCTAGCGTGTATGCTGAGACTCTTTTGGAGCCCTTATGACTTTAGAT 51	
DB	509 TGGAGACATGTGGATACCGAGACGAGACACCATGTTCTGGAGCCCTTCACACATTCGGCT 56	
QY	512 ATGAGCTAGCTGCTCAAGGTACCTTCACAAGGGTTCACAAGTCTCTCAAAATCTAA 57	
DB	569 ACGCGCGCGCTGCCCGGCGAGATGCTCCGAGGTGTACTCCCGCTACTTCAAGAGAGCTGA 62	
QY	572 AGAATTTATACAGAACGACCAATTCATTAATGGAGAACCAAGTATCTCTCC 63	

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 12:01:27 : Search time 54 Seconds
(without alignments)
7564.693 Million cell updates/sec

Title: US-09-857-612A-13
Perfect score: 1332
Sequence: 1 atgaagaagaagaagaaga.....gtaattgtttcggggttaa 1332

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCtUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46.4	3.5	7218	1	US-08-232-463-14 Sequence 14, Appl
2	45	3.4	7218	1	US-08-232-463-14 Sequence 14, Appl
3	37.6	2.8	3581	2	US-08-738-349-1 Sequence 1, Appl
4	36.8	2.8	1776	3	US-08-655-352-10 Sequence 10, Appl
5	36.8	2.8	1776	4	US-09-258-016-10 Sequence 10, Appl
6	36.8	2.8	1776	4	US-09-257-825B-10 Sequence 10, Appl
7	33.8	2.5	3095	6	5231168-1 Patent No. 5231168
8	33.4	2.5	981	4	US-09-134-001C-982 Sequence 881, App
9	32.6	2.4	2402	4	US-09-221-017B-881 Sequence 1, Appl
10	32.4	2.4	5943	1	US-08-206-176-1 Sequence 1, Appl
11	31.6	2.4	2263	2	US-08-687-865A-1 Sequence 1, Appl
12	31.6	2.4	2263	4	US-09-043-711-1 Sequence 1, Appl
13	31.6	2.4	50000	4	US-09-146-053-4 Sequence 4, Appl
14	31.4	2.4	2845	1	US-08-289-653-1 Sequence 1, Appl
15	31.4	2.4	4822	4	US-09-484-970B-106 Sequence 106, App
16	31.4	2.4	5785	3	US-08-480-640A-221 Sequence 221, App
17	31.4	2.4	5785	4	US-08-686-968C-221 Sequence 221, App
18	31.4	2.4	5785	4	US-08-488-237A-221 Sequence 221, App
19	31.2	2.3	1026	4	US-09-134-001C-660 Sequence 660, App
20	31.2	2.3	1960	4	US-09-280-116-133 Sequence 133, App
21	31.2	2.3	3763	1	US-07-792-865D-1 Sequence 1, Appl
22	30.8	2.3	35828	4	US-09-449-218D-17 Sequence 17, Appl
23	30.6	2.3	1866	4	US-08-936-165A-202 Sequence 202, App
24	30.6	2.3	1950	4	US-09-134-001C-1054 Sequence 1054, App
25	30.6	2.3	16593	4	US-08-961-527-52 Sequence 52, Appl
26	30.6	2.3	30549	4	US-09-134-001C-322 Sequence 322, App
27	30.4	2.3	3541	4	US-09-180-439-5 Sequence 5, Appl

28	30.4	2.3	3979	4	US-09-180-439-1 Sequence 1, Appl
29	30.4	2.3	3979	4	US-09-180-439-2 Sequence 2, Appl
30	30.4	2.3	4123	4	US-09-180-439-7 Sequence 7, Appl
31	30.4	2.3	4746	2	US-08-819-288-2 Sequence 2, Appl
32	30.4	2.3	4746	4	US-09-400-348-2 Sequence 2, Appl
33	30.4	2.3	4747	4	US-08-261-822A-2 Sequence 2, Appl
34	30.4	2.3	4747	5	PCT-US95-07744A-2 Sequence 2, Appl
35	30.4	2.3	6042	4	US-08-261-822A-1 Sequence 1, Appl
36	30.4	2.3	6042	5	PCT-US95-07744A-1 Sequence 1, Appl
37	30.4	2.3	6172	2	US-08-819-288-1 Sequence 1, Appl
38	30.4	2.3	6172	4	US-09-400-348-1 Sequence 1, Appl
39	30.2	2.3	369	4	US-09-134-001C-177 Sequence 177, App
40	30.2	2.3	452	4	US-09-221-017B-146 Sequence 146, App
41	30.2	2.3	732	4	US-09-134-001C-124 Sequence 124, App
42	30.2	2.3	1958	2	US-08-390-888A-1 Sequence 1, Appl
43	30	2.3	2877	4	US-09-619-353-1 Sequence 1, Appl
44	29.8	2.2	1563	4	US-09-177-249-3 Sequence 3, Appl
45	29.8	2.2	3222	4	US-08-936-165A-111 Sequence 111, App

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHERFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
CLASSIFICATION: 435
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-Fls
US-08-232-463-14
Query Match 3.5%, Score 46.4, DB 1, Length 7218;

Best Local Similarity 6.28; Pred. No. 0.0002;
Matches 23; Conservative 192; Mismatches 153; Indels 0; Gaps 0;

```

Oy 909 TTTTCAGCTCATGACATGTTGATTTTCTAAAGACATTGTTTCTCGAAGGGGTTTA 968
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1479 TTACCTATCTATGCAAGTACTTAAGAGATAGAGAAATTGGACRNRNRNRNRNR 1420
Oy 969 TCCTATGAACAGCAATCTACCTTGTATAGGAACATAAAGACCAACACTGCTTAT 1028
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1419 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1360
Oy 1029 AACTTTATATAGGGAACGGGAGTGGGAACCTTGGAAACATTTTGTATGGAAGTGA 1088
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1359 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1300
Oy 1089 TTTTGTGAACGGCCAGAAATATCATATGGGATGGGATGGAACGGTGAACCTGTGAG 1148
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1299 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1240
Oy 1149 CTTGTGGCGCTTCAATCATATGAAAGAGAGAAATCAATACCTTAAAGTGTAA 1208
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1239 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1180
Oy 1209 GATAGTGGGCTGTCTCATCTACTTAAAGATGGAAGTGCCTAATAATAGT 1268
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1179 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1120
Oy 1269 AGGTGAGA 1276
    : : : : :
Db 1119 RRRRRRRR 1112

```

RESULT 2

```

US-08-232-463-14
: Sequence 14, Application US/08232463
: Patent No. 5670367
: GENERAL INFORMATION:
: APPLICANT: DORNER, F.
: APPLICANT: SCHEIFLINGER, F.
: APPLICANT: FALKNER, F. G.
: TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 1800 Diagonal Road, Suite 500
: CITY: Alexandria
: STATE: VA
: COUNTRY: USA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/232.463
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/935.313
: FILING DATE:
: APPLICATION NUMBER: EP 91 114 300.6
: FILING DATE: 26-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29, 768
: REFERENCE/DOCKET NUMBER: 30472/114 IMMU
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300
: TELEFAX: (703)683-4109
: TELEX: 899149
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:

```

```

: LENGTH: 7218 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: CLONE: pTZgpt-FLS
US-08-232-463-14

```

Query Match 3.4%; Score 45; DB 1; Length 7218;
Best Local Similarity 5.8%; Pred. No. 0.00057;
Matches 24; Conservative 212; Mismatches 177; Indels 0; Gaps 0;

```

Oy 180 CTCCTACTTCATCTCGAATCATGATGACCTCATCAAGAAAGAAATGATGTTGAG 239
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1067 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1126
Oy 240 ACTTTGGTTGATTCAGAGTCACTACTGCTCTTCACTCAATGTTGGCCGAACGAT 299
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1127 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1186
Oy 300 GACCTTCATTTACACCAAGAACTGATGATTAATCTCAACACTGCTGGGTTGAGACCG 359
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1187 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1246
Oy 360 GGTCCCTCACTTGTGTCGACCAACTCTCTCTATCTCAATCCTGTCGAAGCATAT 419
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1247 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1306
Oy 420 CACCGGATACATGCGACCCCTGTGATGATTCATTACAAAGCTTGGTACGCTGATGTA 479
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1307 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1366
Oy 480 GACTCTGTTGGAGCCCTTATGACTTTAGATATGCTAGCTGGAAGTCACCCCTTC 539
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1367 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1426
Oy 540 ACAAGTGGTTCCAAGTCTCCTCAAGATTAAGAAATTTGATAGAGAACGAA 592
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1427 YYYYYYGTACCAAAATCTCTATCTCTTTAACTTTCATTAAGATAGTAA 1479

```

RESULT 3

```

US-08-738-349-1/C
: Sequence 1, Application US/08738349
: Patent No. 5869638
: GENERAL INFORMATION:
: APPLICANT: Takeshita, Sunao
: APPLICANT: Okazaki, Makoto
: APPLICANT: Kawai, Shinji
: APPLICANT: Tsujimura, Atsushi
: APPLICANT: Amano, Egon
: TITLE OF INVENTION: Bone-Related Cadherin-like Protein and
: TITLE OF INVENTION: Process for its Production
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/738.349
: FILING DATE: 25-OCT-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/364.439

```

FILING DATE:
APPLICATION NUMBER: US 08/112,061
FILING DATE: 26-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Barker, M. P.
REGISTRATION NUMBER: 32,013
REFERENCE/DOCKET NUMBER: 02481.1323-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3581 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Mus musculus
STRAIN: osteoblastic cell line MC3T3E1
FEATURE:
NAME/KEY: CDS
LOCATION: 284..2671
US-08-738-349-1

Query Match
Best Local Similarity 50.6%; Score 37.6; DB 2; Length 3581;
Matches 91; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 1129 GGAACGGTGAAGCTTGTGAGCTTGTGGCGCTTCATCACTATGGAAGAGAAAAT 1188
DB 3070 GAACCTGTGCTTCTTAATTTTGTGAGCTTGTGACAAATTAATAATTAAG 3011
QY 1189 CAATACCTTAAGCGTTAGATAGATGGGTCTCATCTTAATTAAGATGAA 1248
DB 3010 AGACAGCTTAGGTGTTCAACACATAGACGATTTCCCTTCATTTAGATTACCGTGCAC 2951
QY 1249 GTTGCACTAAATGAATAGTAGTGAGATTCTCAATTAATTCATGCTGAGCTCGGT 1308
DB 2950 TTCCTCTTCAATTAATATTTTGTGTCCAGAGTGCTTTTCTTCTACTGAAGTTGAT 2891

RESULT 4
US-08-655-352-10/c
Sequence 10, Application US/08655352
Patent No. 6077991
GENERAL INFORMATION:
APPLICANT: Bachettlira W. Pooviah, Zhihua Liu,
APPLICANT: Shameekumar Patil, Daisuke Takezawa
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
PRODUCTION OF MALE-STERILE PLANTS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klarquist Sparkman Campbell Leigh &
ADDRESSEE: Winston, LLP
STREET: One World Trade Center
STREET: 121 S.W. Salmon Street
STREET: Suite 1600
CITY: Portland
STATE: Oregon
COUNTRY: United States of America
ZIP: 97204
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk, 3-1/2 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/655,352
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,449

FILING DATE: October 14, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Alan E.
REGISTRATION NUMBER: 35,123
REFERENCE/DOCKET NUMBER: 4630-45000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391
TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1776 base pairs
TYPE: nucleic acid
STRANDEDNESS: double stranded
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: Tobacco CcAmk cDNA and deduced amino-acid
FEATURE:
NAME/KEY: protein-coding sequence (not including stop codon)
LOCATION: nucleotides 20-1570
US-08-655-352-10

Query Match
Best Local Similarity 56.7%; Score 36.8; DB 3; Length 1776;
Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1074 TTATGGGAAGCGTATTTGATGACGCCCAATATATCATATGCGGATGATGACAC 1133
DB 1703 TTAATTTAAAGACCTTTTTCAGTAGATGATTAATAATATGATTTGTCATAGANA 1644
QY 1134 GTGACACTGGTGAAGCTTGTGGCGCTTCATCACTATGAAAGGAGAAAATCAATA 1193
DB 1643 GTTAGACATTAATAGATAGATGCGGTATTATTAATAAAGGCAAAATTCATA 1584

RESULT 5
US-09-258-016-10/c
Sequence 10, Application US/09258016
Patent No. 6362395
GENERAL INFORMATION:
APPLICANT: Bachettlira W. Pooviah, Zhihua Liu,
APPLICANT: Shameekumar Patil, Daisuke Takezawa
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
PRODUCTION OF MALE-STERILE PLANTS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klarquist Sparkman Campbell Leigh &
ADDRESSEE: Winston, LLP
STREET: One World Trade Center
STREET: 121 S.W. Salmon Street
STREET: Suite 1600
CITY: Portland
STATE: Oregon
COUNTRY: United States of America
ZIP: 97204
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk, 3-1/2 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/258,016
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Stephens Jr., Donald L.
REGISTRATION NUMBER: 34,022
REFERENCE/DOCKET NUMBER: 4630-51994
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391
TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 1776 base pairs
TYPE: nucleic acid
STRANDEDNESS: double stranded
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
DESCRIPTION: Tobacco Ccalk cDNA and deduced amino-acid
DESCRIPTION: sequence
FEATURE:
NAME/KEY: protein-coding sequence (not including
NAME/KEY: stop codon)
LOCATION: nucleotides 20-1570
US-09-258-016-10

Query Match 2.8%; Score 36.8; DB 4; Length 1776;
Best Local Similarity 56.7%; Pred. No. 0.11;
Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 1074 TTATGGGAAGTGATTTTGATGACGCCAGAAATATCATATGCGGATGTGATGGAAC 1133
DB 1703 TTAATTTTAAAGCATCTTTTACGAGTAGATTAATAATATGATTCATGAGAA 1644
QY 1134 GGTGAAGTGGTGGCTTGTGGCGCTTCATCATCTATGGAAGAGCAAAATCAATA 1193
DB 1643 GTTAGAGATATTAAATGAATTAGCGTGTATTACAAATTAAGAGCAAAATTCATA 1584

RESULT 6
US-09-257-825B-10/C
Sequence 10, Application US/09257825B
Patent No. 6403352
GENERAL INFORMATION:
APPLICANT: Poovaihan, Bachettira W.
APPLICANT: Patil, Shameekumar
APPLICANT: Takezawa, Daisuke
TITLE OF INVENTION: Compositions and Methods for Production of Male-Sterile Plants
FILE REFERENCE: 4630-51993
CURRENT APPLICATION NUMBER: US/09/257,825B
CURRENT FILING DATE: 1999-02-25
PRIOR APPLICATION NUMBER: US 08/655,352
PRIOR FILING DATE: 1996-05-23
PRIOR APPLICATION NUMBER: US 60/014,743
PRIOR FILING DATE: 1996-03-28
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn version 3.1
SEQ ID NO 10
LENGTH: 1776
TYPE: DNA
ORGANISM: Nicotiana tabacum
US-09-257-825B-10

Query Match 2.8%; Score 36.8; DB 4; Length 1776;
Best Local Similarity 56.7%; Pred. No. 0.11;
Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 1074 TTATGGGAAGTGATTTTGATGACGCCAGAAATATCATATGCGGATGTGATGGAAC 1133
DB 1703 TTAATTTTAAAGCATCTTTTACGAGTAGATTAATAATATGATTCATGAGAA 1644
QY 1134 GGTGAAGTGGTGGCTTGTGGCGCTTCATCATCTATGGAAGAGCAAAATCAATA 1193
DB 1643 GTTAGAGATATTAAATGAATTAGCGTGTATTACAAATTAAGAGCAAAATTCATA 1584

RESULT 7
5231168-1
Patent No. 5231168
APPLICANT: DIEGIEL, MORTEN; BORRE, MARTIN; JEPSEN, SOREN;
VUUST, JENS; RIENECK, KLAUS; WIND, ANNETTE; JAKOBSEN, PALLE H.
TITLE OF INVENTION: MALARIA ANTIGEN
NUMBER OF SEQUENCES: 19
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/409,658

FILING DATE: 18-SEP-1989
SEQ ID NO: 1
LENGTH: 3095
5231168-1

Query Match 2.5%; Score 33.8; DB 6; Length 3095;
Best Local Similarity 48.7%; Pred. No. 1.5;
Matches 92; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
QY 1082 AAGGTGATTTTGATGAAGCCAGCAAAATATCATATGCGGATGTGATGAGCGTGAAC 1141
DB 1199 AACATGAATAGTAGAGTTGAAGAAATTTCTACCGAAGATTAATAATGAAAAGTCAAC 1258
QY 1142 TGGTGAAGTGTGGCGCTTCAATCATCTATGGAAGAGCAAAATCAATCAATTAAG 1201
DB 1259 ATGAATAGTAGAGGTGAAGAAATTTCTACCGAAGATGATTAATAATGAAAAGTGAAC 1318
QY 1202 TGGTTAAGATAGATGGCGTCTCATCTACTTCAATCTTAAAGATGAGTGCACATAATG 1261
DB 1319 ATGAATAGTAGAGGTGAAGAAATTTCTACCGAAGATTAATAATGAAAAGTCAACATG 1378
QY 1262 AATAGTAG 1270
DB 1379 AATAGTAG 1387

RESULT 8
US-09-134-001C-982
Sequence 982, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucelte-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: CTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 982
LENGTH: 981
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-982

Query Match 2.5%; Score 33.4; DB 4; Length 981;
Best Local Similarity 51.0%; Pred. No. 0.98;
Matches 79; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 390 TCTCATCTCATCTCTCGTCTCAACATATACCGGATATACGACCCCTGTGATTC 449
DB 291 TTTTAACCTTATCTCGGATTAACATATATCTCCATATATTGAAGACAGTAGACA 350
QY 450 ATTACAAAAGCTTGGCTAGCGTATGATGAGACCTGTGTTGGAGCCCTTATGACTTAG 509
DB 351 AATGATATATGATGAATTAAGAAGACGATACATGTGTTCTGGCGCTCATATTCTAG 410
QY 510 ATATGCTAGCTGCTGAAGGTCAACCTTCACAG 544
DB 411 TTTTCTGTGGCTCTTATGATCAACGTCCTCAAG 445

RESULT 9
US-09-221-017B-881
Sequence 881, Application US/09221017B
Patent No. 6444799
GENERAL INFORMATION:
APPLICANT: Ross, Bruce C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120

CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
APPLICATION DATA: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Montoy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 881:
SEQUENCE CHARACTERISTICS:
LENGTH: 2402 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...2402
US-09-221-017B-881

Query Match 2.4%; Score 32.6; DB 4; Length 2402;
Best Local Similarity 51.0%; Pred. No. 3.1;
Matches 77; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 857 ACCCAAAATTTTGGCTCCAAAACCAATAGATGATACCTCCATTAAGGCTTATTTAG 916
DB 662 ATCGAATATCATGTTTATTTCCCGAGCAGAGAGCTGACCATTTTGGCTTTGCG 721
QY 917 CTCATGACATGGTTGATTTTCTAAAGACATTTGTTTCTGAAGGGTTTATCCTTATG 976
DB 722 TTGGTGCACATCGGTTTGTGTGATGATGCTTATCCGCGACAGGCTTTTATGGAG 781
QY 977 AAACAGCAATTTACCTTGATAGGAAACAT 1007
DB 782 ATACAGGAAGTCTCATTGGAGCAATCAT 812

RESULT 10
US-08-206-176-1
; Sequence 1, Application US/08206176
; Patent No. 5639940

GENERAL INFORMATION:
APPLICANT: Garner, Ian
APPLICANT: Dalrymple, Michael A
APPLICANT: Prunkard, Donna E
APPLICANT: Foster, Donald C
TITLE OF INVENTION: Production of fibrinogen in transgenic
TITLE OF INVENTION: Animals
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zymogenetics, Inc.
STREET: 4225 Roosevelt Way, N.E.
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98105
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/206,176
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Parker, Gary E
REGISTRATION NUMBER: 31-648
REFERENCE/DOCKET NUMBER: 93-15
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-548-2329
TELEFAX: 206-547-8080 ext 322
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5943 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: Human fibrinogen A-alpha chain
FEATURE:
NAME/KEY: CDS
LOCATION: join(31..84, 1154..1279, 1739..1922, 3055..3200,
LOCATION: 3786..5210)
US-08-206-176-1

Query Match 2.4%; Score 32.4; DB 1; Length 5943;
Best Local Similarity 48.4%; Pred. No. 6.2;
Matches 90; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 861 AAAAATTTTGGCTCTCAAAAACCAATAGTATGATACCTCAATTAAGGCTTATTCAGCTCA 920
DB 2657 AAATGATTTGAATGCAATCAACCAACATTTTAACTCCAAATTAACCTTTAAATTC 2716
QY 921 TGACATGTTGATTTTCTAAAGACATGTTTCCGTAAGGGTTTATTCCTTATGAAC 980
DB 2717 AAAGAAATTTCTCTTCTATATTTCTTTGGCATTAATGCTATTATGAGACATTTAAC 2776
QY 981 AGAATTTACCTTGATAGGAAACATAAAGCACACCAAGTGCCTATTAAGTTGATTTAT 1040
DB 2777 TGGCATTCATGGAAGGCTGCAAGGCAATTAATTCAAAGTCAAAATGCCCCATAGGTT 2836
QY 1041 GGGAAC 1046
DB 2837 TTGAAC 2842

RESULT 11
US-08-687-865A-1
; Sequence 1, Application US/08687865A
; Patent No. 5955596
; GENERAL INFORMATION:
; APPLICANT: Jones, Kevin F.

```

? APPLICANT: Zagursky, Robert J.
? APPLICANT: Ool, Peggy
? TITLE OF INVENTION: The Nuca Protein of Haemophilus
? TITLE OF INVENTION: Influenzae and the Gene Encoding That Protein
? NUMBER OF SEQUENCES: 23
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: American Cyanamid Company
? STREET: One Cyanamid Plaza
? CITY: Wayne
? STATE: New Jersey
? COUNTRY: U.S.A.
? ZIP: 07470
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/687,865A
? FILING DATE: 26-JUL-1996
? CLASSIFICATION: 536
? ATTORNEY/AGENT INFORMATION:
? NAME: Gordon, Alan M.
? REGISTRATION NUMBER: 30,637
? REFERENCE/DOCKET NUMBER: 33,250-00
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 201-831-3244
? TELEFAX: 201-831-3305
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2263 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 229..2037
? US-08-687-865A-1

```

```

Query Match          2.4%: Score 31.6: DB 2: Length 2263:
Best Local Similarity 50.7%: Pred. No. 6.3:
Matches 76: Conservative 0: Mismatches 74: Indels 0: Gaps 0:
QY 1067 CATTTGTTTATGGAAGCGTATTTGATGAAAGCGCCAGAAATATCATATGGGATGTCG 1126
DB 1694 CATTCCTTATGTCGAGGTATTCGTATGAAAGCGAATGAAACCAATGCGGAAGTGA 1753
QY 1127 ATGGAACGCTGAACCTTGTCGACCTTGGCGCTTCAATCACTACTATGGAAGAAGAGAAA 1186
DB 1754 ACCGTTTACTGAGTGTGGAAGTCTTGAAATTAACCAACCAACAATGGAACCGATTGATG 1813
QY 1187 ATCAATACCTTAAGCTGTTAAGATGATG 1216
DB 1814 ATACAAACGTTATCTTGTGGAACAATG 1843

```

```

RESULT 12
US-09-043-711-1
? Sequence 1, Application US/09043711
? Patent No. 6221365
? GENERAL INFORMATION:
? APPLICANT: Jones, Kevin F.
? APPLICANT: Zagursky, Robert J.
? APPLICANT: Ool, Peggy
? TITLE OF INVENTION: The Nuca Protein of Haemophilus
? TITLE OF INVENTION: Influenzae and the Gene Encoding That Protein
? NUMBER OF SEQUENCES: 23
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: American Cyanamid Company
? STREET: One Cyanamid Plaza
? CITY: Wayne
? STATE: New Jersey

```

```

? COUNTRY: U.S.A.
? ZIP: 07470
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/043,711
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/687,865
? FILING DATE: 26-JUL-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Gordon, Alan M.
? REGISTRATION NUMBER: 30,637
? REFERENCE/DOCKET NUMBER: 33,250-00
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 201-831-3244
? TELEFAX: 201-831-3305
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2263 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 229..2037
? US-09-043-711-1

```

```

Query Match          2.4%: Score 31.6: DB 4: Length 2263:
Best Local Similarity 50.7%: Pred. No. 6.3:
Matches 76: Conservative 0: Mismatches 74: Indels 0: Gaps 0:
QY 1067 CATTTGTTTATGGAAGCGTATTTGATGAAAGCGCCAGAAATATCATATGGGATGTCG 1126
DB 1694 CATTCCTTATGTCGAGGTATTCGTATGAAAGCGAATGAAACCAATGCGGAAGTGA 1753
QY 1127 ATGGAACGCTGAACCTTGTCGACCTTGGCGCTTCAATCACTACTATGGAAGAAGAGAAA 1186
DB 1754 ACCGTTTACTGAGTGTGGAAGTCTTGAAATTAACCAACCAACAATGGAACCGATTGATG 1813
QY 1187 ATCAATACCTTAAGCTGTTAAGATGATG 1216
DB 1814 ATACAAACGTTATCTTGTGGAACAATG 1843

```

```

RESULT 13
US-09-146-053-4
? Sequence 4, Application US/09146053A
? Patent No. 639349
? GENERAL INFORMATION:
? APPLICANT: Ryan, James W.
? APPLICANT: Sprinkle, Jerry Joe Curtis
? APPLICANT: Venema, Richard C.
? TITLE OF INVENTION: Human Amino-peptidase P Gene
? FILE REFERENCE: MCG103
? CURRENT APPLICATION NUMBER: US/09/146,053A
? EARLIER FILING DATE: 1998-09-02
? EARLIER APPLICATION NUMBER: 60/057,854
? EARLIER FILING DATE: 1997-09-02
? NUMBER OF SEQ ID NOS: 7
? SOFTWARE: Patentin Ver. 2.0
? SEQ ID NO 4
? LENGTH: 50000
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-146-053-4

```

Query Match 2.4%: Score 31.6: DB 4: Length 50000;

Db 2252 TGATACACTCATGCGCTCTTCATATGCGGGGGTTTCAGGCTCATGTTTATGCTTC 2193
Qy 277 A 277
1
Db 2192 A 2192

Search completed: February 15, 2003, 13:16:59
Job time : 116 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 12:58:02 : Search time 66 Seconds
(Without alignments)
10279.077 Million cell updates/sec

Title: US-09-857-612A-13
Perfect score: 1332
Sequence: 1 atgaagaagaacaaga.....gtaattgttcggggttaa 1332

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 424239 seqs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database: Published_Applications_NA:*

1: /cgn2_6/pdata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/pdata/1/pubpna/PCr_NEW_PUB.seq:*
3: /cgn2_6/pdata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/pdata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/pdata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/pdata/1/pubpna/PCrUS_PUBCOMB.seq:*
7: /cgn2_6/pdata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/pdata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/pdata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/pdata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/pdata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/pdata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/pdata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/pdata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	436.6	32.8	1242	9	US-09-938-842A-376
2	155.6	11.7	1048	9	US-09-998-059-8
3	155.6	11.7	1059	9	US-09-998-059-20
4	155.6	11.7	1080	9	US-09-998-059-27
5	64.6	4.8	1744	10	US-09-919-497-30
6	64.6	4.8	1744	10	US-09-880-107-2245
7	60.4	4.5	2680	9	US-09-978-295A-156
8	60.4	4.5	2680	9	US-09-978-697-156
9	60.4	4.5	2680	9	US-09-978-192A-156
10	60.4	4.5	2680	9	US-09-999-832A-156
11	60.4	4.5	2680	9	US-09-978-189-156
12	55.4	4.2	248	10	US-09-923-876-2515
13	51.2	3.8	2180	9	US-10-001-054-7
14	36	2.7	954	10	US-09-947-971-3
15	36	2.7	640681	10	US-09-790-988-1
16	34.2	2.6	312	10	US-09-880-107-3898
17	34.2	2.6	2364	12	US-10-044-090-37
18	34.2	2.6	32199	10	US-09-764-855-210
19	33.8	2.5	239	10	US-09-983-965-5347

20	33.6	2.5	1154	10	US-09-867-550-1491	Sequence 1491, Ap
21	33.4	2.5	878	9	US-09-828-995B-57	Sequence 57, Appl
22	33.4	2.5	878	9	US-09-828-995B-59	Sequence 59, Appl
23	33.4	2.5	1095	9	US-09-828-995B-65	Sequence 65, Appl
24	33.4	2.5	1095	9	US-09-828-995B-67	Sequence 67, Appl
25	33.4	2.5	1158	9	US-09-828-995B-63	Sequence 63, Appl
26	33.4	2.5	1158	9	US-09-828-995B-64	Sequence 64, Appl
27	33.4	2.5	1454	9	US-09-828-995B-60	Sequence 60, Appl
28	33.4	2.5	1454	9	US-09-828-995B-62	Sequence 62, Appl
29	33.2	2.5	698	10	US-09-823-101-3	Sequence 3, Appl
30	33.2	2.5	1722	10	US-09-815-742-7160	Sequence 7160, Ap
31	33	2.5	1998	10	US-09-833-381-1139	Sequence 1139, Ap
32	32.6	2.4	2769	10	US-09-739-254-53	Sequence 53, Appl
33	32.6	2.4	2769	10	US-09-904-615-53	Sequence 53, Appl
34	32.6	2.4	3849	10	US-09-764-864-560	Sequence 560, Ap
35	32.6	2.4	3850	10	US-09-764-864-113	Sequence 113, Ap
36	32.6	2.4	6819	9	US-10-098-841-4	Sequence 4, Appl
37	32.2	2.4	484	9	US-10-046-935-620	Sequence 620, Ap
38	32.2	2.4	484	9	US-09-878-178-620	Sequence 620, Ap
39	32.2	2.4	1412	10	US-09-925-301-568	Sequence 568, Ap
40	32.2	2.4	1969	10	US-09-729-674-7	Sequence 7, Appl
41	32	2.4	267	10	US-09-923-876-4976	Sequence 4976, Ap
42	32	2.4	2082	9	US-09-893-519A-105	Sequence 105, Ap
43	32	2.4	6114	9	US-09-938-842A-2700	Sequence 2700, Ap
44	32	2.4	6804	10	US-09-740-041-3	Sequence 3, Appl
45	31.8	2.4	765	10	US-09-897-214-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-938-842A-376
Sequence 376, Application US/09938842A
Patent No. US0020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPT300-3
CURRENT APPLICATION NUMBER: US/09/938, 842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227, 866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264, 647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300, 111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 376
LENGTH: 1242
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-376
Query Match 32.8%; Score 436.6; DB 9; Length 1242;
Best Local Similarity 62.3%; Pred. No. 1.3e-119;
Matches 767; Conservativity 0; Mismatches 434; Indels 30; Gaps 4;
OY 74 TGTCTTCTTATGCACTGTGGGCAACACCTTGACCTTGTATTAACAGTA 133
DB 2 TGACCTCATGTGTCAACCTGTGGGTACCAACGTTTACCTGTGATCTGCAGGAA 61
OY 134 ACGGAGGACCAATATAGACAGGTGACCAACCTTACTTCACTTCACT 193
DB 62 ACGGAGGTACAGGTATAGAGTACGCTGACAGCATATACAGCAAGTGTCTGCT 121
OY 194 GCGAATCATGTACCTCTCATCAAGAAAGA--ATGATGTGTTCAGACTTGGTTG 250
DB 122 GTAGACGTGTGTATATCCGATTTATAGAGAGTGTGTGTTAGGCTATGCTTGC 181

QY	251	ATTCGAGTGATACCTGGCTCCCTTTCACCTAAAGCTTGGCGAAGCATACCCCTTCAT	31
Db	182	ATCGACGAGGTATATATGCTCCCTTCACGAGTGCTTCACCGATTCGATATGTTGACT	241
QY	311	ACCAACAAGAACTGCATGATTAATTCAACAACCTCCGTGGGTGAGACCCGGTCCCTACT	370
Db	242	ATGACCCCTGATTTGGATGATTAACCAAAATGCTCCTGGTGTCCAAACCCGGGTTCATAT	301
QY	371	TTGGTTCACCAACGCTCTCTCTATCTCAATCTCTGCTCAAGCAATACCGGATPACA	430
Db	302	TGCGTTGGACCAAAATCACTCTTATACCTCGAACCTGCTCTCGAGATGCGCACATCTTACA	361
QY	431	TGGCAACCCCTGGTAGATTTCATTAC - - AAAAGCTTGGCTACGCTGATGAGACACTGT	487
Db	362	TGGAACATTTGGTGAAGGCTCTACAGAAAAAATGCGGGTATGTTTAAGACAAACCATCC	421
QY	488	TTTGAGGCCCCCTATATGACTTTTATGATATGCTAGCTGCTGAAGGTCAACCTTCACAAGTG	547
Db	422	TAGAGAGCTCCATATGATTTTACAGTAGCGGCTGCTGCTGGGCCCAACCGTCCCGTGTAG	481
QY	548	GTTCCCAAGTTCTCAAGATCTAAAGAAATTTGATAGAAAGAACGACCAATTCCAATAAG	607
Db	482	CCTCACAAGTTCCTACAGACTTCAAAACAATTTGGTGGAAAACTAGCAGCGAAGCAAG	541
QY	608	GGAAGCCAGTGATACCTTCTCTCCACAGTTTAAAGAGCCTATTGTCTTCAACACTATA	667
Db	542	GAAAGCCAGTGATACCTCTCTCCCATAGCCCTAAGAGACATTTTCGTCCTCATATTCCTCA	601
QY	668	ATPAAACACCCCTCTTGGGGCAAAAAATTCATCAAAACCTTCATGCTCTTTCAGCTC	727
Db	602	ACCGTACCACCCCTTCATGCGCGCGGCAAGTACATCAACACTTTGTGCACTCGCGCC	661
QY	728	CATGGGGTGCTGATATAGAGAAATGTACACCTTGCATCTGGCAACACTTTGGGAGTGC	787
Db	662	CATGGGGTGAGCATCTCTCAGATGAAGACATTTGGTCTGGCAACACTCGGTGCTCC	721
QY	788	CCCTATAGTGACCCCTTATATAGTAGGATATAAAGAAAGCTCCGAGACTTAACCTTTGGC	847
Db	722	CTTTAGTTAACCCCTTCTCTGTCAGACGCGCATCAGAGGACCTCCGAGAGTAAACAATGGC	781
QY	848	TTTTGGCCCAACCCAAAAATTTTTT - - -GGTCTCAAAAACCAATAGATACCTCAATTA	904
Db	782	TACTTTCATCTACCAAGGTGTTTCACGACACGAAGCTAAACCCGTTGTCTGTAACCTCCACG	841
QY	905	GCGCTTATTCAGCTCATGACATGGTGTATTTCTAAAGACATTTGGTTCCTGAAGGG	964
Db	842	TTTAACCTACACGCTTACGAGATGGATCGGTTTTCAGACATTTGGATTTCTCAACAAGAG	901
QY	965	TTTATCTTATGAAACGGAATTCACCCCTGATAGGGAACATTAAGAACCAACCAAGTGC	1024
Db	902	TTTGTGCTTACACAGACAGGTGTTGCTTTTACAGAGGAGCTGATGACTCCGGGGTGC	961
QY	1025	CTATTAACCTGATTTATGGAAGGAGGAGTGGAACCTTGGAAACATTTGTTATGGGAAG	1084
Db	962	CAGTACACTTGATATATGGAAGAGAGGATGATACACCGGAGTTTGTATGATGGAAGA	1021
QY	1085	GTGATTTTGATGAAGCGCCCAGAATATCATATGGGATGCTGATGGAAGCGTGAACCTGG	1144
Db	1022	GAGGATTCGATTAAGCAACAGAAATTAAGATATGGAATGAGATGGAGCGGTTAATTTGG	1081
QY	1145	TGAGCTTGTGGCGCTTCATCTACATATGMAAGAGAGAAAAATCAATACCTTAAAGTGG	1204
Db	1082	CGAGCTTAGCAGCTT - - - - -GAAAGTCATTAAGCTTGAACACCG	1122
QY	1205	TTTAAGATAGATGGGGTCTCATACCTCAATACCTTAAAGATGAAGTGGACATTAATGAA	1266
Db	1121	TAGAGATGATGGAGTTCGCAATCTATCTTAAGAGCAGATCGCATTTAAAGAGA	1180
QY	1265	TATAGGTGAGATTAATCTTCAATTAATTTCA 1295	
Db	1181	TTATGAAGCAGATTTCAATTAATTAATTA 1211	

```

RESULT 2
US-09-998-059-8/c
; Sequence 8, Application US/09998059
; Publication No. US20030005485A1
;
GENERAL INFORMATION:
;
APPLICANT: Onitroge, John B.
APPLICANT: Benning, Christoph
APPLICANT: Gao, Hongbo
APPLICANT: Girke, Thomas
APPLICANT: White, Joseph A.
; TITLE OF INVENTION: Plant Seed Specific Promoters
;
FILE REFERENCE: MSU-06689
;
CURRENT APPLICATION NUMBER: US/09/998,059
;
CURRENT FILING DATE: 2001-11-30
;
PRIOR APPLICATION NUMBER: 60/250,401
;
PRIOR FILING DATE: 2000-12-01
;
NUMBER OF SEQ ID NOS: 60
;
SOFTWARE: PatentIn version 3.1
;
SEQ ID NO 8
;
LENGTH: 1048
;
TYPE: DNA
;
ORGANISM: Arabidopsis thaliana
;
US-09-998-059-8

```

Query Match	11.7%;	Score 155.6;	DB 9;	Length 1048;
Best Local Similarity	65.7%;	Pred. NO. 3.9e-36;		
Matches 243; Conservative	0;	Mismatches 124;	Indels 3;	Gaps 1

QY	45	ACTACAGTTACACATAGTGTGGTACGTGCTATGTCATGTGCATGTGGGACACAA	104
Db	639	AGTCATAGCCGATACTCGTTGTGGTGACGATGAGCTCGATGTGTCAACTGTGGGTACAA	580
QY	105	CCTCGACCCCTTGTGATCTAAATACCAAGTAACGGAGGACCAACTAGAACCAAGTTGAC	164
Db	579	CGTGTACCCCTTGTGATCTGTTGCTGGTTCAGAAAGGAGGTAAACAGCTAGAGGTACGGCTGA	520
QY	165	CAATGACATACAAAGCCCTCTACTTTCATCTCGAATCATGTGATCCCTCTCATCAAGAAAA	224
Db	519	CAGAGAAATCAAAAGCCAGTAGTGCTGTGGTATAGCAGCTGGTTATATCGAATTATAAGAA	460
QY	225	GA---ATGAGTGGTTCAGACTTTGGTGTGATTCAGTGCATACCTTCTCTCTTCACTCA	281
Db	459	GAGTGGTGGATGTTTATAGGCTATGATGTTGCATGCAAGCAGTGTATTGTCTCTCCCTTACACAG	400
QY	282	ATGCTTTGCCGAAGCATGACCCCTTCATTAACCAACCAAGAACTCGATGATTACTTCAACAC	344
Db	399	GTCGTTCAAGGATGCGAATGATGTTGTACTATGAACCCGATTTTGGATGATTACCAAAATGC	340
QY	342	TTCCTGGGGTTGAGACCCGGGTCCCTCACTTGGTTTCCACCAACTCTTCTTATCTCAA	401
Db	339	TTCCTGGTGTCAAAACCCGGGGTCTCTCATTTTCGTTGCACCAAAATCATCTTATACCTCGA	280
QY	402	TTCGTCGCTC 411	
Db	279	CCCTCGCTC 270	

RESULT 3
 US-09-998-059-20/c
 ; Sequence 20, Application US/09998059
 ; Publication No. US20030005485A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ohlrogge, John B.
 ; APPLICANT: Benning, Christoph
 ; APPLICANT: Gao, Hongbo
 ; APPLICANT: Girke, Thomas
 ; APPLICANT: White, Joseph A.
 ; TITLE OF INVENTION: Plant Seed Specific Promoters
 ; FILE REFERENCE: MSU-06689
 ; CURRENT APPLICATION NUMBER: US/09/998,059
 ; CURRENT FILING DATE: 2001-11-30
 ; PRIOR APPLICATION NUMBER: 60/250,401

OY	580	ATACAGAAACAACCAATTCCATAAATGGGAGCACAGTGTACTTCCTGCCACAGTTTA	639
Db	831	GGAGGACATGCACG-----CTGCCATATGGGAACCCTGTCTTCTCTCATTTGGCCACACCTTC	885
OY	640	GGAGGCGCTATTTGTCTCAACACTACTAATAATGAAAACCCCCCTCTTGGCGCAGAAAAATTTC	699
Db	886	GCGCTGTACACTTGCTCTATTTCTCTGCTCGCCACACCCCAGGCGTGGAAAGACCGCTTT	945
OY	700	ATCAAACACTTCATTCGTCTCTTCAAGCTCCATGGGGGTGGTGTCTTATACAGAANAATGTACACC	759
Db	946	ATTATATGGCTTCATCTCTCTTTGGGGCTCCCTGGGGGTGGCTCCATCAAGCCCATGCTGGTCC	1001
OY	760	TTTTSCATCTGGCAACACTTTGGGAGTAGTCCCCCTAGTGGAACCTTTATTAGTGAAGGATGAA	819
Db	1006	TTTGCCCTCAGGTGCACACAGGGGCAATCCCATCATGCTCCAGCATCAAGCTGAAGAAGGAG	1066
OY	820	CAAGGAAGCTCCGAGATTAACCTTTGGCTTTTGCC	854
Db	1066	CAGCGCATTAACCAACACCTCCCTCCCTGGATGATTTTCC	1100

```

RESULT 6
US-09-880-107-2245
: Sequence 2245, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: APPLICANT: Gene Logic, Inc.
: TITLE OR INVENTION: Gene Expression Profiles In Liver Cancer
: FILE REFERENCE: 44921-5028-WO
: CURRENT APPLICATION NUMBER: US/09/880,107
: CURRENT FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2245
: LENGTH: 1744
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 MI2625
US-09-880-107-2245

```

	Query Match	4.8%	Score 64.6:	DB 10:	Length 1744;	
	Best Local Similarity	49.0%:	Pred. No. 5.7e-09:			
	Matches	311:	Conservative	0:	Mismatches 294;	Indels 30; Gaps 4
Oy	220	AAAAAGAAATGCATGGTTCACACTTTGGTTTGATTCCAGTGTCACTACTGCTCCTTTCACT	279			
Db	496	AAGACAGAGGACCTTCTCACCATCTGGCGTGGATCTCAACATGTCTTACCCTTGGGGATA	555			
Oy	280	CATAGCTTTGGCGAAGCAGCATGCCCTTCATTATCCACAAGAAGCTCGATGATTACTTCAC	339			
Db	556	GACTGCTGCATTCGATTAACACCAAGSGTTTGTCATCAACACCGAAGCTCTGGCTCGTCCAC	615			
Oy	340	ACTCTGGGGGTTGAGACCCGGGTCCCTCACTTTGGTTGCCAACACTCTTCTCTATCTC	399			
Db	616	GCCTCTGGTGTCCAAATCCGGCTCCCTGSCCTTTGGCAAAGACTACTCTGTGAAGTACCG	675			
Oy	400	AATCCTCTCTCAAGCATATCACCGGATACATGAGCACCCCTGATGATTATTACAAAG	459			
Db	676	GACAG-----CAGCAACTGGCAGGGATACCTGCACACACTGTGTGAAGAACTGTCAC	729			
Oy	460	CTTGGCTACCGCTGATGTGTGAGACTCTGTTTGGAGCCCCCTTAGACTTTTAGATATAGTCTTA	519			
Db	730	AATGGCATCGTGGCGGACAGACTGTGGCGCGCCCTTATGACTG-----	776			

QY	520	GCTCTAAGGTACCCCTTACAAAGGGGTTCCAACTCCCTCAAGATTTAAAGAAATTTG	579
Db	777	GGCGCTGAGACCGCGCCAGCAGAGGAGT-----ACTACCGCAAGCTGCAGCGCTGCT	830
QY	580	ATAGAAAGACAAAGCAATTTCAATTAATGGGAACCGCATATCTCTCCACAGTTTA	639
Db	831	GAGAGCAATGCAACG-----CTGCGCTATGGGAACCGCTGTCTTCTCATTTGGCACAGCTTC	885
QY	640	GGAGGCGTATTTGTCTTCAACAATTAATAATAGAAACCCCCCTTTGGCGCAAAAAATTC	699
Db	886	GGCTGTCTACACTTGTCTCTATTTCTGTCGTGCGCACGCCACCGCTCGGAAGACCGCTTT	945
QY	700	ATCAAAACACTTTCATTTGCTCTTTACGCTCCATGGGGGTGGTGTATATGACCAATTTGACAC	759
Db	946	ATTGATGGCTTCAATCTCTCTTTGGGGCTCCCTCGGGGTGGCTCCATCAAGCCCATGCTGGTC	1005
QY	760	TTTTCATCTGGCAACACTTTGGGAGTGGCCCTTATGTGACACCTTTATATGAGGATGAA	819
Db	1006	TTTGCGCTACAGTGTGCAACCGAGGCGATCCCATCATTTCCAGATCAAGCTGCAAAAGAGAG	1065
QY	820	CAAGAAGAGCTCGGAGATTAACTTTGGCTTTTGGC	854
Db	1066	CAGCGCATTAACCAACACCTCCCTCGATGTTTTC	1100

```

: RESULT 7
: US-09-978-295A-156
: Sequence 156, Application US/09978295A
: Patent No. US20020156006A1
: GENERAL INFORMATION:
: APPLICANT: Asinkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Baton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerlitsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kiljavan, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James;
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P2630P1C11
: CURRENT APPLICATION NUMBER: US/09/978,295A
: CURRENT FILING DATE: 2001-10-15
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450

```


; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085580
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085573
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085704
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085697

Query Match 4.5% Score 60.4; DB 9; Length 2680;
 Best Local Similarity 44.3% Pred 1.3e-07;
 Matches 418; Conservative 0; Mismatches 496; Indels 30; Gaps 3;

QY 205 TACCTCTCATAGAAAAAGATGGATGTTGAGACTTGTGGATTGATTCAGTGTATATA 264
 DB 207 TACCTCTCTCAAGAAAGACCAAACTACTTACAAATCGGTGTAACCTGGAACTGTG 266
 QY 265 CTTCCTCTTTCATCATATGCTTTGGCGAAGCATGACCTTTCATTACCAACAAGACTC 324
 DB 267 CTGCGTGTATCATGTAGCTGTGATGACAAATATCAGCTGGTTTACAACAAACATCC 326
 QY 325 GATGATTAATCAACACTCTGGGGTGGAGACCGGGTCCCTCACTTGTGGTCCACCAAC 384
 DB 327 AGGGCCACCCATTTCTCTATATGTGTGTGATGATGATGATGATGATGATGATGATG 386
 QY 385 TCTCTTCTATCTCAATCTCTGTCTCAAGATATCACCGGATACAGGACCCCTGTGTA 444
 DB 387 TCACGTGAGTTCTCTGAGACCCCAAGAAAGACGCTGGGTTCTTATTTCCACACATG 446
 QY 445 GATTCATTACAAAGCTTGGCTACGCTGATGTGAGACTGTGTGGAGCCCTTATGAC 504
 DB 447 GAGACCCCTTGTGGCTGAGGCTACACAGCGGTGAGGATGTCGAGGGGCTCCATATGAC 506
 QY 505 TTTAGATATGCTCTAGCTGCTGAGGTCACCTTCAAGTGGGTTCCAAAGTTCTCAAA 564
 DB 507 TGGGCCCGAGCCCAATATGAAACGGGCCCTTCTGCGCCCTCCGAGATGATGACG 566
 QY 565 GATCTAAAGATTTGATATAGAGACCAAGCAATTCATATGGAAGCCAGTATACTT 624
 DB 567 GAGATGTACACAGCTGT-----ATGGGGGGCCCGCTGGTGCTG 602
 QY 625 CTCTCCCAAGTTTAGAGGCGCTATTTGCTTCAACTACTTAATATGAAGCCCGCTCT 684
 DB 603 GTTCCCAAGATGTGGGCAACATGTACACGCTTCTTCTGACCGGAGCGGAGGCG 662
 QY 685 TGGCCCAAAAAATTCATCAAACTTCAATGCTTTCAGCTGCAATGGGGTGTCTATA 744
 DB 663 TGGAGGAAATATATATCCGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 722
 QY 745 GACGAATGTACACTTGTGATCTGGCAACACTTTGGGAGTGCCCTAGTGGACCCCTT 804
 DB 723 AAGACCTCTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782
 QY 805 TTAGTGGAGGATGAGCAAAAGAGCTCCGAGATACCTTTGGCTTTTGGCTTAACCCAAA 864
 DB 783 AAGATCCCGGAGACACGCGGTCACTCTCTCCACACCTGCTCTCCCTCAACATAC 842
 QY 865 ATTTTGTCTCTCAAAAAACAAATAGTATGATCTCAATTAAGCCCTTATTCAGACTATGAC 924
 DB 843 AACTGTGTACCTGAGAAAGTGTGTGAGACACCAATCACTACACACTGCGGGGAC 902
 QY 925 ATGTTGATTTTCTTAAAGACATGTTTCTGTAAGGGGTTTATCTTATGAAACAGCA 984
 DB 903 TACCGCAAGTTCTTCCAGACATCGGCTT---TGAAGATGGCTGCTCATGCGGAGAGAC 959
 QY 985 ATTCTACCTTGATAGGAAACATTAAGACACCAAGTCCCTATATACCTTGTATTTATGGA 1044
 DB 960 ACAGAAAGGCTGTGTGGAAGCAAGATGCGACCTGCGAGCTGCGCTCTATGCT 1019
 QY 1045 ACGGAGATGGGAACCTTGGAAACATGTTTATGGAAGGATGATTTGATGAAAGGGA 1104
 DB 1020 ACTGGCGTCCCAACACAGACTCTTCTTACTATGAGA--GCTTCCCTGACCGTGACCT 1076

QY 1105 GAATATCATATGGGATGTGATGGAACGGTGAACCTTGTGAG 1148
 DB 1077 AAAATCTGCTTGTGATGACGCGCATGCTGTAAGTGAAGAG 1120

RESULT 8
 US-09-978-697-156
 ; Sequence 156, Application US/09978697
 ; Patent No. US20020169284A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Baker Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Ferrara, Napoleon
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gottsen, Mary E.
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth J.
 ; APPLICANT: Klyavin, Ivar J.
 ; APPLICANT: Kuo, Sophia S.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Shelton, David L.
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: P2630P1C27
 ; CURRENT APPLICATION NUMBER: US/09/978,697
 ; CURRENT FILING DATE: 2001-10-16
 ; PRIOR APPLICATION NUMBER: 09/918585
 ; PRIOR FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/064249
 ; PRIOR FILING DATE: 1997-11-03
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066364
 ; PRIOR FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: 60/077450
 ; PRIOR FILING DATE: 1998-03-10
 ; PRIOR APPLICATION NUMBER: 60/077632
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077641
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077649
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077791
 ; PRIOR FILING DATE: 1998-03-12
 ; PRIOR APPLICATION NUMBER: 60/078004
 ; PRIOR FILING DATE: 1998-03-13
 ; PRIOR APPLICATION NUMBER: 60/078886
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078936
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078939
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/079294

...
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079669
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28

PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 4.5%; Score 60.4; DB 9; Length 2680;
Best Local Similarity 44.3%; Pred. No. 1.3e-07;
Matches 418; Conservative 0; Mismatches 496; Indels 30; Gaps 3;

QY 205 TACCTTCATCAAGAAAAGATGATGCTTACAGACTTTGGTTGATTCAGTCTATA 264
DB 207 TACCTTCATCAAGAAAAGATGATGCTTACAGACTTTGGTTGATTCAGTCTATA 266
QY 265 CTTCCTCTTTCATCAATGCTTTGCCGACGACGACCTTCATTTACCAACCAAGACTC 324
DB 267 CTGCCTGATCATTTGACTGCTGATTTGACAATATCAGGCTGTTTACCAACCAACATTC 326

PRIOR FILING DATE:	1998-03-31
PRIOR APPLICATION NUMBER:	60/080165
PRIOR FILING DATE:	1998-03-31
PRIOR APPLICATION NUMBER:	60/080194
PRIOR FILING DATE:	1998-03-31
PRIOR APPLICATION NUMBER:	60/080327
PRIOR FILING DATE:	1998-04-01
PRIOR APPLICATION NUMBER:	60/080328
PRIOR FILING DATE:	1998-04-01
PRIOR APPLICATION NUMBER:	60/080333
PRIOR FILING DATE:	1998-04-01
PRIOR APPLICATION NUMBER:	60/080334
PRIOR FILING DATE:	1998-04-01
PRIOR APPLICATION NUMBER:	60/081070
PRIOR FILING DATE:	1998-04-08
PRIOR APPLICATION NUMBER:	60/081049
PRIOR FILING DATE:	1998-04-08
PRIOR APPLICATION NUMBER:	60/081071
PRIOR FILING DATE:	1998-04-08
PRIOR APPLICATION NUMBER:	60/081195
PRIOR FILING DATE:	1998-04-08
PRIOR APPLICATION NUMBER:	60/081203
PRIOR FILING DATE:	1998-04-09
PRIOR APPLICATION NUMBER:	60/081229
PRIOR FILING DATE:	1998-04-09
PRIOR APPLICATION NUMBER:	60/081955
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/081817
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/081838
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/082568
PRIOR FILING DATE:	1998-04-21
PRIOR APPLICATION NUMBER:	60/082569
PRIOR FILING DATE:	1998-04-21
PRIOR APPLICATION NUMBER:	60/082704
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION NUMBER:	60/082797
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION NUMBER:	60/082766
PRIOR FILING DATE:	1998-04-23
PRIOR APPLICATION NUMBER:	60/083366
PRIOR FILING DATE:	1998-04-27
PRIOR APPLICATION NUMBER:	60/083322
PRIOR FILING DATE:	1998-04-28
PRIOR APPLICATION NUMBER:	60/083392
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/083495
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/083496
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/083499
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/083545
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/083554
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/083559
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/083500
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/083742
PRIOR FILING DATE:	1998-04-30

PRIOR APPLICATION NUMBER:	60/084366
PRIOR FILING DATE:	1998-05-05
PRIOR APPLICATION NUMBER:	60/084414
PRIOR FILING DATE:	1998-05-06
PRIOR APPLICATION NUMBER:	60/084441
PRIOR FILING DATE:	1998-05-06
PRIOR APPLICATION NUMBER:	60/084633
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/084633
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/084640
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/084640
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/084652
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/084652
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/084653
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/084653
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/085338
PRIOR FILING DATE:	1998-05-13
PRIOR APPLICATION NUMBER:	60/085338
PRIOR FILING DATE:	1998-05-13
PRIOR APPLICATION NUMBER:	60/085323
PRIOR FILING DATE:	1998-05-13
PRIOR APPLICATION NUMBER:	60/085587
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/085700
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/085688
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/085579
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/085580
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/085577
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/085704
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/085697
PRIOR FILING DATE:	1998-05-15

Query March	4.5%;	Score 60.4;	DB 9;	Length 2680;
Best Local Similarity	44.3%;	Pred. No. 1.3e-07;		
Matches 418; Conservative	0;	Mismatches 496;	Indels 30;	Gaps 3.

QY	205	TACCCCTCATTCAAGAAAGAAAGATGATGATGCTGACAGCTTTGGTTGGTTCCAGTGCATATA	266
Db	207	TACCTCTGCTCCAAAGAACCGAAGACTACTTCCAACTCTGCGCTGGAACCTGGAACCTGCTG	266
QY	265	CTTGCTCCTTTCACTCAATGCTTTTGCCGAAGCATGACCCCTTCATTTACACCAAGAACTC	324
Db	267	CTGCGCTGCATTCATTCGATGCTGGATTGACAAATATCAGGGCTGGTTTACACAAACATATCC	324
QY	325	GATGATTACTTCAACACTCCTGGGGTTGAGACCCGGGTCCCTCACTTTGGTTCCACCAAC	386
Db	327	AGGGCCACCCAGTTTCCCTGATGGGTGTGATGTACGTCTCCCTGGCTTTGGGAAGACCTTC	386
QY	385	TCTCTTCTATCTCATTCGATCCGCTCTCAAGATATTCACCCGGATATTCATGAGCACCCCTGTATA	444
Db	387	TCACTGAGATCTCCGAGACCCACCAAGACAGCTGGTTCCTATTTCACACCATATGTTG	444
QY	445	GATTCATTACAAAGACTTGGCTACGCTGATGATGATGATGAGACTGTTTGGAGGCCCTTATGAC	504
Db	447	GAGAGCCTTGTGGGCTGTGGGCTACACACGGGTGATGAGATGTCCGAGGGGCTCCCTATGAC	504
QY	505	TTTAGATATGCTAGCTGCTGAAGGTCAACCCCTTCAACAAGTGGGTTCCAAAGTTCTCAAA	564
Db	507	TGGCGCCGAGCCCCCAATGAAAGAGGGGCCCTACTTCTGTGCCCTCCGCGGAGATGATGAG	566
QY	565	GATTTAAGATTTTGATAGGAAGAAAGCAAGCAATTCCAATATATGGGAAGCCAGTATACTT	624
Db	567	GAGATGTACACGCTGT-----ATGGGGGCCCCCGTGGTCTCG	602

;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081898
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07

;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 4.5%; Score 60.4; DB 9; Length 2680;
Best Local Similarity 44.3%; Pred. No. 1.3e-07;
Matches 418; Conservative 0; Mismatches 496; Indels 30; Gaps 3;

QY 205 TACCCTCATCAGAAAAAGATGATGCTTCAGACTTTGGTTGATTCACAGTCTATA 264
DB 207 TACCCTCTCCCAAGAAACCGAAAGCTACATTCAGATTCGAGAACTGAACTGTG 266
QY 265 CTGCTCCTTCACCTCAATGCTTTCCGACGATGACCTTCATTCACCAAGAACTG 324
DB 267 CTGCTGTCATCATTCATGCTGCTGGAATATACATATACAGCTGCTTTACAAACAAACATCC 326
QY 325 GATGATTCCTCAACACTCTGCGGGTTGAGACCGGGTCCCTCACTTTGGTTCCACCAAC 384
DB 327 AGGGCCACCCAGTTTCTCGATGCTGTGATGATGATGATGATGATGATGATGATGATG 386
QY 385 TCTCTCTCTATCTCAATCTCTCTCTCAAGCATATACCGGATACATGACACCTCGTGA 444
DB 387 TCACGTGAGTTCTCTGACCCCAAGAAAGCAGCGGTGCTCTATTTCCACACCATGTG 446
QY 445 GATTCATTACAAAAGCTTGGCTAGCTGATGATGATGATGATGATGATGATGATGATGATG 504
DB 447 GAGAGCTTTGGGCTGGGCTGACACAGGGGTGAGATGTCGAGGGGCTCCATATAC 506
QY 505 TTTAGATATGCTAGCTGTAAGGTCACCTTCACAAAGTGGGTTCCAAAGTTCCTCAAA 564
DB 507 TGGCGCCGAGCCCAATGAAGAAAGCGGCCCTACTTCTGCGCCCTCGGAGATGATGAG 566
QY 565 GATCTAAAGATTGATGAAAGAACAAAGCAATTCATTAATGGAAACCACTGATCTT 624
DB 567 GAGATGTACCAAGCTGT-----ATGGGGGGCCCGGTGTGTG 602
QY 625 CTCTCCACAGTTTGAAGAGGCTATTTGCTCAACTACTATAATAGAAACCCCTCTCT 684
DB 603 GTTCCCCACAGTATGGCAACATGTACACGCTCTACTTTGACAGGCGACGCGAGGCC 662
QY 685 TGGCGCAAAAATTCATCAACACTTCATTCCTTCAGCTTCATTCATTCATTCATTCATTCAT 744
DB 663 TGAAGAGCAAGTATATCCGGGCTTGTGACATGGGTGGGCGGCTGTGGCC 722
QY 745 GACGAATGTACACTTTGCTGCAACACTTTGGAGAGTGGCCCTTATGTTGACCTTTTA 804
DB 723 AAGACCCCTGCGCTCTGCTTCAGAGAACAAACCGATTCACAGTATGCGGCGCTGTG 782
QY 805 TTAGTGAGGATGAACAAGAGCTCCAGAGTAACTTTGCTTCCTTCCTTACCAACCAAA 864
DB 783 AAGATCCGGGAGCAGCAGCGGTCACTGCTTCACACCTGCTGCTGCTTACAACTAC 842
QY 865 ATTTTGTCTCAAAAACCAATAGTATACCTCAATTTAGGCTTATTCAGCTATGAC 924
DB 843 ACATGTACCTGAGAAAGTGTGTGTGACAGACCCCAATCACTACACACTGGGGAC 902

QY 925 ATGCTTGAATTTCTTAAAGACATGTTGTTTCTGTAAGGGTTTATCTTATGAACAGCA 984
DB 903 TACCCGACATTTCTTCCAGACATCGGCTT---TGAAGATGGCTGGCTCATGCGGACGAC 959
QY 985 ATTTCTACCTTGTATAGCAATTAAGACACACAGTGCCTTAACTGTATATAGGA 1044
DB 960 ACAGAAAGGCTGTGTGAACCCAGATGCCAGCTGGCGGTGACGACTGCTCTATGCT 1019
QY 1045 ACCGAGTGGAGACCTTGGAAACATTTGTTATGGAAGGTGATTTGTGAACGCGCA 1104
DB 1020 ACTGGGCTCCACACACACATCCTTCTACTATGAGA---GCTTCCTGACCGTACCCCT 1076
QY 1105 GAAATATCATATGAGGATGTGTATGGAACGCTGACTGTGTAG 1148
DB 1077 AAAATCTGCTTGTGTGACGCGCATGTACTGTGAAG 1120

RESULT 11
US-09-978-189-156
; Sequence 156, Application US/09978189
; Publication No. US20030004102A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Fliviaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumes, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630P1C7
; CURRENT APPLICATION NUMBER: US/09/978,189
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704

1	PRIOR FILING DATE: 1998-04-22	2	PRIOR APPLICATION NUMBER: 60/082804
3	PRIOR FILING DATE: 1998-04-22	4	PRIOR APPLICATION NUMBER: 60/082700
5	PRIOR FILING DATE: 1998-04-22	6	PRIOR APPLICATION NUMBER: 60/082797
7	PRIOR FILING DATE: 1998-04-22	8	PRIOR APPLICATION NUMBER: 60/082766
9	PRIOR FILING DATE: 1998-04-23	10	PRIOR APPLICATION NUMBER: 60/083336
11	PRIOR FILING DATE: 1998-04-27	12	PRIOR APPLICATION NUMBER: 60/083322
13	PRIOR FILING DATE: 1998-04-28	14	PRIOR APPLICATION NUMBER: 60/083392
15	PRIOR FILING DATE: 1998-04-29	16	PRIOR APPLICATION NUMBER: 60/083495
17	PRIOR FILING DATE: 1998-04-29	18	PRIOR APPLICATION NUMBER: 60/083496
19	PRIOR FILING DATE: 1998-04-29	20	PRIOR APPLICATION NUMBER: 60/083499
21	PRIOR FILING DATE: 1998-04-29	22	PRIOR APPLICATION NUMBER: 60/083545
23	PRIOR FILING DATE: 1998-04-29	24	PRIOR APPLICATION NUMBER: 60/083544
25	PRIOR FILING DATE: 1998-04-29	26	PRIOR APPLICATION NUMBER: 60/083558
27	PRIOR FILING DATE: 1998-04-29	28	PRIOR APPLICATION NUMBER: 60/083559
29	PRIOR FILING DATE: 1998-04-29	30	PRIOR APPLICATION NUMBER: 60/083500
31	PRIOR FILING DATE: 1998-04-29	32	PRIOR APPLICATION NUMBER: 60/083722
33	PRIOR FILING DATE: 1998-04-30	34	PRIOR APPLICATION NUMBER: 60/084336
35	PRIOR FILING DATE: 1998-05-05	36	PRIOR APPLICATION NUMBER: 60/084414
37	PRIOR FILING DATE: 1998-05-06	38	PRIOR APPLICATION NUMBER: 60/084411
39	PRIOR FILING DATE: 1998-05-06	40	PRIOR APPLICATION NUMBER: 60/084637
41	PRIOR FILING DATE: 1998-05-07	42	PRIOR APPLICATION NUMBER: 60/084639
43	PRIOR FILING DATE: 1998-05-07	44	PRIOR APPLICATION NUMBER: 60/084640
45	PRIOR FILING DATE: 1998-05-07	46	PRIOR APPLICATION NUMBER: 60/084598
47	PRIOR FILING DATE: 1998-05-07	48	PRIOR APPLICATION NUMBER: 60/084600
49	PRIOR FILING DATE: 1998-05-07	50	PRIOR APPLICATION NUMBER: 60/084627
51	PRIOR FILING DATE: 1998-05-07	52	PRIOR APPLICATION NUMBER: 60/084643
53	PRIOR FILING DATE: 1998-05-07	54	PRIOR APPLICATION NUMBER: 60/085339
55	PRIOR FILING DATE: 1998-05-13	56	PRIOR APPLICATION NUMBER: 60/085338
57	PRIOR FILING DATE: 1998-05-13	58	PRIOR APPLICATION NUMBER: 60/085323
59	PRIOR FILING DATE: 1998-05-13	60	PRIOR APPLICATION NUMBER: 60/085582
61	PRIOR FILING DATE: 1998-05-15	62	PRIOR APPLICATION NUMBER: 60/085700
63	PRIOR FILING DATE: 1998-05-15	64	PRIOR APPLICATION NUMBER: 60/085580
65	PRIOR FILING DATE: 1998-05-15	66	PRIOR APPLICATION NUMBER: 60/085737
67	PRIOR FILING DATE: 1998-05-15	68	PRIOR APPLICATION NUMBER: 60/085704
69	PRIOR FILING DATE: 1998-05-15	70	PRIOR APPLICATION NUMBER: 1998-05-15

; PRIOR APPLICATION NUMBER: 60/085697

Query Match	4.58;	Score 60.4;	DB 9;	Length 2680;
Best Local Similarity	44.38;	Pred. No. 1.3e-07;		
Matches 418;	Conservative 0;	Mismatches 496;	Indels 30;	Gaps 3

OY	205	TACCTCTCATCAAGAAAAGAAATGGAATGGATGGTCCAGACCTTGGTTTGATTCACAGTGTCA	264
OY	206		
Db	207	TACCTCTGCTCCAGAAAGAACCCGAACCTCTTCTCACATCTGGCTCAACCTGGAACTGCTG	266
OY	265	CTTCTCTCTTTCACCTCACTCAATCTTGGCCGAAGCATGACCTCTTCAATTCACCAAGAACTC	324
OY	266		
Db	267	CTGGCTCTGCATCATGTACTGCTGGATTTGAACAAATACAGGCTGTTTACAAACAAACATCC	326
OY	325	GATCATTTACTTCAACACTCTCTGGGGTTTGAGACCCGGGTCCTCTCACTTTGGTTCCACCAAC	384
Db	326		
OY	327	AGGGCCACCCAGTTTCTGTATGGTGTGGATGTACGTGTCCTCCGGGCTTGGGAAGACCTTC	386
OY	385	TCTCTTCTATCTCATCTCTGCTGCTCAAGCATATCACGGATACATGGACACCCCTGGTA	444
Db	387	TCACTGGAGTCTCTGGACCCCAAGAAAGACGGCTGGTTCTCTATTTTCCACACCATGGTGTG	446
OY	445	GATTCATTTACAAAAGCTTGGCTACGGCTGATGTGAGACTCTGTTTGAAGCCCTCTATGAC	504
Db	447	GAGAGCCTTGTGGGCTGTGGGCTCTACACAGGGGTGAGAGATGTCCGAAGGGGCTCCTATGAC	506
OY	505	TTTAGATATGTCTAGCTGCTGAAGGTCAACCTTCACAAGTGGGTTCCAGTTCTCTCAA	564
Db	507	TGGGGCCGAGGCCCAATATGAAAACGGGGCCCTTACTTCTGGCCCTCCGAGATGTATCGAG	566
OY	565	GATCTAAAGAATTGTATAGAAAGAACAGCAATTCATATATGGGAAGCAGATGTACTT	624
Db	567	GAGATGTACCAAGCTGT-----ATGGGGGGCCCGTGATGCTGTG	602
OY	625	CTTCCACAGTTTATGAGAGGCGCTATTGTGTCCTCAACTACTAAATAGAAACCCCTCTCT	684
Db	603	GTTCGCCACATATAGGCGAACATGTACAGCTCTACTTCTGTACAGGAGCGAGCCGAGGCC	662
OY	685	TGGCGCAAAAATTCATCAAAACACTTCATTTGCTCTTTCAGCTCCATCGGGGTGTCTATA	744
Db	663	TGGAAAGCAAGTATATCCGGGCTCTGTGTCAGTGGGTGGCCCTCGGGGGGGCGGTGACC	722
OY	745	GACCAAAATGTACACCTTTGCATCTGCGACACACTTGGGAGTGGCCCTAGTGGACCTTTA	804
Db	723	AAGACCTCGCGCTCTCGGCTTACGAGAGAACACACCGGATCCCAAGTCATCGGGGCCCTG	782
OY	805	TTAGTAGAGGATGTGAACAAAGACCTCGAGAGTAACTTTGGCTTTTGGCTTAACCAAA	864
Db	783	AAGATCCGGGAGCGACGCGGTCAAGCTGTCTCCACAGCTGGCTGTGCTGCCATCAACATAC	842
OY	865	ATTTTGTGCTCAAAAACCAATAGTATTAATCACTCAATTAGGCCCTTATTCAGCTCATGAC	924
Db	843	ACATGTGCACCTGAGAGGTGTGTGTCAGACACCCCAATCAACTACACTGTGGGAGC	902
OY	925	ATGGTTGATTTTCTAAAAGCATTTGGTTTTCCITGAAGGGGTTTATCCTTATTAACACAGA	984
Db	903	TACCGCAAGTCTTCCGAGCAATCGGCTT--TGAAGATGTGGCTGTATGCGGAGGAC	959
OY	985	ATTCTACCTCTTGATAGGAGACATAAAAGCCACCAAGTGCCTCTATACTGTATTTATGGGA	1044
Db	960	ACAGAAAGGCTGTGGTGAAGCAGACATGCCACCTGGGCTGACGTGCATGCTCTATAGT	1019
OY	1045	ACGGGAGTGGGACCTTGGAAACATTTGTTTATGGGAAAAGGTATTTTGATGAACGGCCA	1104
Db	1020	ACTGGCGTCCCAACACAGACTCCTTCTACTATGAGA---GCTCCCTATGACCGTACCT	1076
OY	1105	GAATATCATATGGGGATGCTGATGCAAGCGTAACTTGGTAG	1148
Db	1077	AAAATCTGCTTTGGTGAACGGCGATGTAGCTGTAACTTAAAG	1120

RESULT 12
US-09-923-876-2515

```
; Sequence 2515, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Laigudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ico)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 2515
; LENGTH: 248
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700160724H1
US-09-923-876-2515

Query Match          4.2%; Score 55.4; DB 10; Length 248;
Best Local Similarity 54.7%; Pred. No. 1.1e-06;
Matches 110; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 423 CGGATACATGGCACCCCTGTATATTCATTACAAAGCCTTGCTAGCGCTGATGTGAGAC 482
DB 26 CTGCTGCTCTCGGAAGCTGAGCCCGCACGGAAGACATGGGATACCGACGAGACGAC 85
QY 483 TCTGTTGAGCCCTTATGACTTTAGATGTTGCTAGCTGCTGAAGGTCACCTTCACA 542
DB 86 CATGTTGAGCCCTTACACTTCCTGCTAGCGCGCCCTCCCGGCTGACGACGACGAC 145
QY 543 AGTGGCTTCCAGCTTCTCAAGATCTMAAGAAATTTGATGAAGAGCAAGATTCCAA 602
DB 146 GGTGTACTCCGCTACTTCAGAGCTGATGAGTGTGCTGAGCGGCAAGAGAGAGAC 205
QY 603 TATGGGAAGCCAGTGATCT 623
DB 206 CCGGAAGAGCGCTCATCTCT 226

RESULT 13
US-10-001-054-7
; Sequence 7, Application US/10001054
; Publication No. US20020192209A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Baker, Kevin
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin
; APPLICANT: Hebert, Carolyn
; APPLICANT: Henzel, William
; APPLICANT: Kabakoff, Rhona
; APPLICANT: Shelton, David
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC
; FILE REFERENCE: P303481PCT
; CURRENT APPLICATION NUMBER: US/10/001,054
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
```

```
; PRIOR APPLICATION NUMBER: 60/082999
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/090691
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/096891
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 60/096894
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100263
; PRIOR FILING DATE: 1998-09-14
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/101476
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/107783
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/108849
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 60/112420
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/113296
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/115554
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/116533
; PRIOR FILING DATE: 1999-01-20
; PRIOR APPLICATION NUMBER: 60/123618
; PRIOR FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: 60/131294
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/162506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/187202
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/209832
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: 60/232887
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/180997
; PRIOR FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: 09/218517
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 09/284291
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380138
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380913
; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: 09/403297
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 09/423741
```

PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/866034
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/882636
PRIOR FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/924419
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/927796
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/929404
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 09/941992
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: 09/946374
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: PCT/US98/18824
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: PCT/US99/00106
PRIOR FILING DATE: 1999-01-05
PRIOR APPLICATION NUMBER: PCT/US99/05028
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: PCT/US99/08615
PRIOR FILING DATE: 1999-04-20
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/20111
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00376
PRIOR FILING DATE: 2000-01-06
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04341
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: PCT/US00/04342
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/06884
PRIOR FILING DATE: 2000-03-15
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/13705
PRIOR FILING DATE: 2000-05-17
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/22031
PRIOR FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: PCT/US00/23522
PRIOR FILING DATE: 2000-08-23

PRIOR APPLICATION NUMBER: PCT/US00/30873
PRIOR FILING DATE: 2000-11-10
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: PCT/US01/06666
PRIOR FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: PCT/US01/17092
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: PCT/US01/17800
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: PCT/US01/21066
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/21735
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: PCT/US01/27099
PRIOR FILING DATE: 2001-08-29
NUMBER OF SEQ ID NOS: 91
SEQ ID NO 7
LENGTH: 2180
TYPE: DNA
ORGANISM: Homo Sapien
US-10-001-054-7

Query Match 3.8%; Score 51.2; DB 9; Length 2180;
Best Local Similarity 46.0%; Pred. No. 6; le-05;
Matches 250; Conservative 0; Mismatches 28; Indels 6; Gaps 2;

605 ATGGAGGCGAGTGATATCTCTCCACAGTTAGAGGCGCTATTGCTACACTAC 664
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83 ATGGGGCCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 142
665 TAAATAGAAACCCGCCCTCTTGGCGCAAAAATTCATCAACACTTCTCTTTCAG 724
143 TGCAGCGGCGACCCGACGCGCGTGAAGGACAGTATATCCGGCGCTGCTGCTGCTG 202
725 CTCATGCGGCTGCTGCTATAGACGAATGTACACCTTTCATCTGCGACACTTGGGAG 784
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
203 CGCCCTGGGGGGCGGTGGCGCAAGACCCCTGCGCTGCTGCTGCTGCTGCTGCTG 262
785 TGCCCCCTAGTGGACCTTATTAGTAGGAGATGAACAAAGAGCTCCGAGAGTAACCTT 844
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
263 TCCAGTCATGCGGCCCTGGAAGATCCGGAGACAGCGGCTCAGCTGCTCCACCACT 322
845 GCGTTTTCCTAACCANAATTTTGGTCCCTCAAAAACCAATAGTGAATCACTCAATTA 904
323 GCGTCTGCTCCCTACACACTACATGCTACCTGAGGAGTGTGCTGACAGACCCACAA 382
905 GGCCTTATTCAGCTCATGACATGCTGATTTCTAANAAGACATGGTTTCCGTAAGGG 964
383 TCAACTACACACTCGGCGGACGCAAGTCTTCCAGACATCGGCTT---TGAAGATG 439
965 TTTATCTTATGAAACAGATTTCTACCTTGATAGGGAACATAAACACACCAAGTGC 1024
440 GCTGCGTCATGCGCGACAGACAGAAAGGCTGGTGAAGCCACGATGCACTGCGCTGC 499
1025 CTATTAATCTGATTTATGAGGAGGAGTGGGACCTTGAANAACATTTGTTATGGGAAG 1084
500 AGCTGCACTGCTCTATGATGCTGCGCTCCACACACGAGACTCTTCTACTATAGAGCT 559
1085 GTGATTTTGTATGAACGCGCAAAATATCATATGCGGATGTGATGAAGAGGTGAACCTGG 1144
560 TCCCTGACCGTGAC---CTTAAATCTGCTTGGTGAAGCGGATGATCTGTGAACCTTGA 616
1145 TGAG 1148
617 AGAG 620

RESULT 14

```
US-09-947-971-3
; Sequence 3, Application US/09947971
; Patent No. US2002014148A1
; GENERAL INFORMATION:
; APPLICANT: LOK, SI
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Human Vomeroneasal Receptor-4
; FILE REFERENCE: 00-83
; CURRENT FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 954
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:2.
; NAME/KEY: misc_feature
; LOCATION: 3, 9, 12, 15, 18, 24, 27, 33, 36, 42, 45, 48, 63, 72, 78,
; LOCATION: 84, 87, 93, 96, 99, 102, 108, 111, 117, 141, 156, 162, 171,
; LOCATION: 177, 186, 189, 201, 204, 213, 216, 219, 222, 228, 234, 237,
; LOCATION: 246, 252, 255, 261, 264, 285, 288, 291, 303, 306, 309
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 318, 333, 339, 342, 345, 357, 360, 363, 369, 372, 375, 387,
; LOCATION: 390, 393, 396, 411, 414, 429, 432, 459, 462, 465, 474, 477,
; LOCATION: 480, 486, 492, 501, 504, 507, 510, 513, 519, 525, 528, 534,
; LOCATION: 537, 540, 543, 549, 552, 564, 567, 570, 573, 582, 591
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 597, 606, 609, 612, 615, 618, 621, 627, 639, 645, 648, 654,
; LOCATION: 660, 666, 675, 678, 681, 684, 693, 699, 708, 714, 720, 726,
; LOCATION: 729, 735, 741, 750, 753, 756, 762, 768, 771, 774, 777, 780,
; LOCATION: 789, 792, 804, 810, 819, 822, 825, 828, 831, 834, 837
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 843, 852, 855, 861, 864, 876, 879, 882, 888, 894, 897, 900,
; LOCATION: 903, 906, 909, 921, 924, 939, 948
; OTHER INFORMATION: n = A,T,C or G
US-09-947-971-3

Query Match 2.7%; Score 36; DB 10; Length 954;
Best Local Similarity 29.8%; Pred. No. 1.3;
Matches 82; Conservative 40; Mismatches 153; Indels 0; Gaps 0;

OY 923 ACATGCTCATTTTCTAAAGACATTGCTTCCCTGAAGGGTTATCTTATGAACAC 982
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 106 ACNNTNTTYYTYYTYYTAAATHTTAAATHTTYYTNCAGAYCARAARMSNAARCCNCAY 165

OY 983 GAATCTACCCCTTAGTAGGACATAAAGCACCAAGTGCCTATTAACCTGTATTATGG 1042
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 166 GAYTNNATHWSTGYAAATVNSCNCNTTATATCAATGCTTATTTTACNGTNGTNCAY 225

OY 1043 GAACGGAGCTGGAAACCTTGAACACATGTTTATGGGAAGGTGATTTTGAATGAACGC 1102
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 226 GCGTGCCGCCNCNGAYATGTGTCARMSNTNCAYTTNGNAAGARTTYAARTTYAARMSN 285

OY 1103 CAGAATATCATATGGGATGCTGATGGAACGCTGAACCTTGTGACCTTGTGGCGCTTC 1162
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 286 YTNWSTNTAYATHAAYMGNTNMNGATGAYTNTGTATHTGYAAYACNTGYTNTYTNMSN 345

OY 1163 AATCACTATGGAAGAGGAAGAAATCAATACCTT 1197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 346 ATHCATCAACGACGCMWSNACNATHMSNCCNMSNAAYTT 380
```

```
RESULT 15
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
```

```
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEKI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIORITY FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 2.7%; Score 36; DB 10; Length 640681;
Best Local Similarity 51.9%; Pred. No. 43;
Matches 81; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

OY 856 AACCCAAAATTTTGGCTGCTCAAAAACCAATAGATACCTCAATTAGGCTTATGCA 915
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 335384 AACCATTAATTAATTTTAACTCTAAATCTTATTAAGAAATCCGATGCTAGTCTTCT 335443

OY 916 GCTCATGACATGCTGATGATTTCTAAAGACATTGTTTCCCTGAAGGGTTATCTTAT 975
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 335444 TTTGATGATCTGTCTATGCTAATTTTCGAAATATTTAGATGAAGTATATGTTAAT 335503

OY 976 GAACACGAATTTCTACCTTGTATAGGGAACATPAAA 1011
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 335504 AAAAAAGAAATTCGTCTCATTTTGTGTCTGAAAA 335539
```

Search completed: February 15, 2003, 14:18:35
Job time : 749 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 11:56:47 : Search time 1677 Seconds
(Without alignments)
12863.678 Million cell updates/sec

Title: US-09-857-612a-13
Perfect score: 1332
Sequence: 1 atgaagaagaagaagaaga.....gtattgtttcggggttaa 1332

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estda:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vtc:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	524.6	39.4	532	13	B1973838
2	515.8	38.7	527	10	AM759200
3	432	32.4	507	14	BM885973
4	333.8	25.1	650	12	BG524903
5	322.6	24.2	673	10	BE322181
6	305.8	23.0	648	14	BQ852109

7	296.2	22.2	630	14	BQ855578
8	256.6	19.3	587	12	BG523648
9	235	17.6	596	12	BG525847
10	234.6	17.6	635	12	BG522086
11	216.8	16.3	616	12	BG526525
12	205.2	15.4	527	14	BQ864610
13	195.4	14.7	608	10	AV441327
14	171.6	12.9	610	10	AV827187
15	170.8	12.8	529	10	AV442635
16	166.8	12.5	724	12	BG127829
17	153	11.5	514	10	AV528103
18	146.6	11.0	537	12	BE355801
19	143.8	10.8	163	13	BE801241
20	143.4	10.8	676	17	BH578169
21	134.8	10.1	401	13	B1139236
22	122.6	9.2	517	13	B1209019
23	122.6	9.2	517	13	B1209572
24	122.6	9.2	517	13	B1210034
25	122.6	9.2	517	13	B1210853
26	117.4	8.8	657	13	B1256228
27	116.6	8.8	593	17	BH777228
28	115.4	8.7	416	13	B1069567
29	114.8	8.6	419	10	AV816862
30	114.8	8.6	731	13	B1406095
31	100.4	7.5	430	10	AV797479
32	91.8	6.9	673	13	B1261764
33	86.8	6.5	637	14	BQ459510
34	85.8	6.4	323	14	BQ862413
35	84.4	6.3	518	12	BF557056
36	79.6	6.0	1350	11	AK004914
37	79.4	6.0	362	14	BQ984301
38	75.8	5.7	669	12	BG836492
39	75.2	5.6	610	13	BM099858
40	74.8	5.6	630	13	BM420443
41	74	5.6	661	13	B1278186
42	73	5.5	524	14	BQ243462
43	73	5.5	640	13	B1283237
44	69.4	5.2	538	10	BE343728
45	69	5.2	377	13	B1268541

ALIGNMENTS

RESULT 1
B1973838
LOCUS
DEFINITION
B1973838 532 bp mRNA linear EST 30-NOV-2001
sa193d08.y1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl065-8607 5' similar to TR:Q9ZM42 Q9ZM42 F17L21.28. ;, mRNA
Sequence.

ACCESSION
B1973838
VERSION
B1973838.1 GI:16348243
KEYWORDS
SOURCE
ORGANISM
soybean.

REFERENCE
AUTHORS
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corvelli,V., Khanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Peterson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

TITLE
JOURNAL
COMMENT

Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available through: Resgen, Invitrogen Corp. 2130
 South Memorial Parkway Huntville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: curesgen.com
 High quality sequence stop: 421.
 Location/Qualifiers
 1..532

FEATURES

Source
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl065-8607"
 /clone_lib="Gm-cl065"
 /tissue_type="germinating shoots"
 /lab_host="DH10B"
 /note="Vector: pBluescript II SK+; Site_1: EcoRI; site_2:
 xhoI; The cDNA library was constructed from mRNA isolated
 germinating shoots of the cultivar Williams. The seeds
 were allowed to germinate for 24 hours prior to being
 cold stressed for 2 days at 4C. Complementary DNA was
 synthesized from mRNA using a primer consisting of a
 poly(dT) sequence with a XhoI restriction site. EcoRI
 adapters were ligated to the blunt-ended cDNA fragments
 followed by XhoI digestion. The cDNA fragments were
 directionally cloned into the EcoRI-XhoI restriction site
 of the pBluescript vector. The ligated cDNA fragments were
 transformed into DH10B host cells (GibcoBRL). This library
 was constructed in the laboratory of Dr. Randy
 Shoemaker."
 BASE COUNT 168 a 87 c 124 g 152 t 1 others
 ORIGIN

Query Match 39.4%; Score 524.6; DB 13; Length 532;
 Best Local Similarity 99.1%; Pred. NO. 3e-148;
 Matches 527; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 759 CTTTCATCTGGCACACTTTGGAGTGGCCCTAGTGGACCTTTATTAGTAGGATGA 818
 |||||||
 Db 1 CTTTCATCTGGCACACTTTGGAGTGGCCCTAGTGGACCTTTATTAGTAGGATGA 60
 |||||||
 OY 819 ACAAGAAGCTCCGAGAGTAACTTTGGCTTTCCTAACCCAAAATTTTGGTCCCA 878
 |||||||
 Db 61 ACAAGAAGCTCCGAGAGTAACTTTGGCTTTCCTAACCCAAAATTTTGGTCCCA 120
 |||||||
 OY 879 AAACCAATAGTGAATCACTCAATTAGGCTTATTCAGCTCATGACATGTTGATTTCT 938
 |||||||
 Db 121 AAACCAATAGTGAATCACTCAATTAGGCTTATTCAGCTCATGACATGTTGATTTCT 180
 |||||||
 OY 999 AAACCAATAGTGGTTTCCGAGAGGTTTATCTTATGAACAGCAATTTCTACCTTGT 998
 |||||||
 Db 181 AAACCAATAGTGGTTTCCGAGAGGTTTATCTTATGAACAGCAATTTCTACCTTGT 240
 |||||||
 OY 999 AGGAACATTAAGACACCAAGTCCCTTAATCTTATTTATGGGAACGGAGTGGGAC 1058
 |||||||
 Db 241 AGGAACATTAAGACACCAAGTCCCTTAATCTTATTTATGGGAACGGAGTGGGAC 300
 |||||||
 OY 1059 CTTGGAACATTTGTTTATGGAAAGGTGATTTGATGAACGGCCAGAAATATCATATGG 1118
 |||||||
 Db 301 CTTGGAACATTTGTTTATGGAAAGGTGATTTGATGAACGGCCAGAAATATCATATGG 360
 |||||||
 OY 1119 GGATGGTATGGAACGGTCACTTGGTGAAGCTTTGGGCTTCAATCACTATGGAAGA 1178
 |||||||
 Db 361 GGATGGTATGGAACGGTCACTTGGTGAAGCTTTGGGCTTCAATCACTATGGAAGA 420
 |||||||
 OY 1179 GGAGAAAATCAATACCTTAAGGTGTTAGATAGATGGGGTGTCTCATACTTCAATACT 1238
 |||||||
 Db 421 GGAGAAAATCAATACCTTAAGGTGTTAGATAGATGGGGTGTCTCATACTTCAATACT 480
 |||||||
 OY 1239 TAAGATCAAGTTCACATAAATAGTAGTAGATTAATCAATTAAT 1290
 |||||||
 Db 481 TAAGATCAAGTTCACATAAATAGTAGTAGATTAATCAATTAAT 532
 |||||||

RESULT 2

AW759200
 LOCUS 527 bp mRNA linear EST 03-DEC-2001
 DEFINITION s138c06.y1 Gm-cl027 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-cl027-3587 5' similar to TR:092M42 092M42 F17L21.28. ;, mRNA
 sequence.
 ACCESSION AW759200
 VERSION AW759200.1 GI:7691066
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 527)
 Shoemaker R., Kelm P., Vodka L., Erpelding J., Corvett V., Khanna
 A., Bolla B., Merritt M., Hillier L., Kneba T., Martin J., Beck C.,
 Wylie T., Underwood K., Steptoe M., Theising B., Allen M., Bowers
 Y., Person B., Swaller T., Gibbons M., Pape D., Harvey N., Schurk
 R., Ritter E., Kohn S., Shih T., Jackson Y., Cardenas M., McCann
 R., Waterston R. and Wilson R.
 Public Soybean EST Project
 Unpublished (1999)
 CONTACT: Shoemaker R./Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available through: Resgen, Invitrogen Corp. 2130
 South Memorial Parkway Huntville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: curesgen.com
 Insert Length: 900 Std Error: 0.00
 High quality sequence stop: 435.
 Location/Qualifiers
 1..527

FEATURES

source

/organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl027-3587"
 /clone_lib="Gm-cl027"
 /tissue_type="cotyledons of 3- and 7-day-old Williams
 seedlings"
 /lab_host="DH10B"
 /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
 xhoI; This cDNA library was constructed from mRNA isolated
 from cotyledons of 3- and 7-day-old Williams seedlings
 which were propagated on paper towels with distilled
 water. The cotyledons were flash-frozen in liquid
 nitrogen, then lyophilized for 72 hours. Unequal amounts
 of mRNA was used for cDNA synthesis. Stratagene's cDNA
 Synthesis Kit (catalog number 200401) was used to
 synthesize the cDNA. First-strand synthesis was
 performed with 5-methyl dCTP, hence the ligated cDNA was
 hemimethylated. A modification of Stratagene's
 first-strand synthesis primer was used. An anchor
 nucleotide (V-A, C, or G) was added to the 3' end of the
 primer [GAGAGAGAGAGAGAGACAGCTGCTGAG(T)18] to anchor
 the primer at the 5' end of the poly(A) tract. After
 second-strand synthesis, the cDNA ends were filled in
 with cloned Pfu DNA, ligated to EcoRI adapters and
 subsequently phosphorylated. The XhoI site within the
 first-strand synthesis primer was then restricted by
 digestion with XhoI; all XhoI sites in the cDNA would be
 protected by their hemimethylated status. The cDNA
 constructs were size-fractionated with a 500 bp cutoff,
 using GibcoBRL Life Technologies' cDNA Size Fractionation
 column. The column eluent was then ligated into
 Stratagene's pBluescript(II) Mx Predigested Vector
 (pBluescript II SK(+)) that has been digested with EcoRI
 and XhoI, and phosphorylated by Stratagene). 97% of the
 white and blue colonies appear to contain recombinant
 plasmids with cDNA inserts, based on size (n=30). This

Library was constructed by Dr. Paul Keim and Dr. Virginia Corryell."

BASE COUNT 156 a 107 c 116 g 148 t

Query Match 38.7%; Score 515.8; DB 10; Length 527;
Best Local Similarity 98.7%; Pared. No. 1.4e-145;
Matches 520; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 626 TTTCCCACTTTAGAGAGGCTATTGTCTTACACTACTATAAGAACCCCTCTT 685
DB 1 TCTCCACAGTTTAGAGGCTATTGTCTTACACTATAAGAACCCCTCTT 60
QY 686 GGGCAAAAAATTCATCAACACTTTCATTCCTTACGCTCCATGGGCTGTATAG 745
DB 61 GGGCAAAAAATTCATCAACACTTTCATTCCTTACGCTCCATGGGCTGTATAG 120
QY 746 ACGAATGTACACCTTGTGATCGCAACACTTGGAGTCCCTAGTGGACCTTTAT 805
DB 121 ACGAATGTACACCTTGTGATCGCAACACTTGGAGTCCCTAGTGGACCTTTAT 180
QY 806 TAGTGAGGATGACCAAGAAGCTCCGAGATACCTTTGGCTTTGGCTAACCCAAAA 865
DB 181 TAGTGAGGATGACCAAGAAGCTCCGAGATACCTTTGGCTTTGGCTAACCCAAAA 240
QY 866 TTTTGGTCTCAAAAACCAATAGTAACTCAATTAGGCTTATTCAGCTCATGACA 925
DB 241 TTTTGGTCTCAAAAACCAATAGTAACTCAATTAGGCTTATTCAGCTCATGACA 300
QY 926 TGGTGTATTTCTTAAAGACATGTTTCTGTAAGGGGTTTATCTTATGAACACGAA 985
DB 301 TGGTGTATTTCTTAAAGACATGTTTCTGTAAGGGGTTTATCTTATGAACACGAA 360
QY 986 TTTACCTTGTATGGGACATTAAGACACACAGTCCCTATTAATCTTGTATGGAA 1045
DB 361 TTTACCTTGTATGGGACATTAAGACACACAGTCCCTATTAATCTTGTATGGAA 420
QY 1046 CGGAGTGGGAACTTGTGAACATTTTATGGAAAGGATTTGATGAACGGCCAG 1105
DB 421 CGGAGTGGGAACTTGTGAACATTTTATGGAAAGGATTTGATGAACGGCCAG 480
QY 1106 AATATCATATGGGATGTGATGAACGTTGATGGAGCTTG 1152
DB 481 AATATCATATGGGATGTGATGAACGTTGATGGAGCTTG 527

RESULT 3
BM885973 507 bp mRNA linear EST 08-MAR-2002
LOCUS BM885973
DEFINITION Gm-c1063-5158 5' similar to TR:Q9ZM42 Q9ZM42 F17121.28.; mRNA
sequence.

ACCESSION BM885973
VERSION BM885973.1 GI:19269717
KEYWORDS soybean.
SOURCE Glycine max.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 507)

REFERENCE
AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corryell, V., Khanna
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Stepoe, M., Theising, B., Allen, M., Bowers
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Wilson, R.
Public Soybean EST Project

TITLE
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:
www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 421.
Location/Qualifiers
1. 507
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1063-5158"
/clone_1bp="Gm-c1063"
/lssue_type="Germinating shoot, 24 hour germination"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from mRNA isolated
from germinating shoots of the cultivar Williams. The
seeds were allowed to germinate for 24 hours prior to
harvesting the germinating shoots. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (Gibco BRL). This
library was constructed in the laboratory of Dr. Randy
Shoemaker at Iowa state university."

FEATURES
source

BASE COUNT 145 a 125 c 100 g 137 t

ORIGIN

Query Match 32.4%; Score 432; DB 14; Length 507;
Best Local Similarity 100.0%; Pared. No. 4.6e-120;
Matches 432; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGAAGGAACAAGAGGCTCAAGATTTAGTGTCTACACTACAGTTACACTA 60
DB 76 ATGAAGAAGGAACAAGAGGCTCAAGATTTAGTGTCTACACTACAGTTACACTA 135
QY 61 GTTGTGTGATGCTGCTCAATTCATGACATGTGGGCAACACTCGACCTTTGATT 120
DB 136 GTTGTGTGATGCTGCTCAATTCATGACATGTGGGCAACACTCGACCTTTGATT 195
QY 121 CTAATACAGGTAACGAGGAGAACCACTAGAACGAGTTGACCAATCAGTACAGCC 180
DB 196 CTAATACAGGTAACGAGGAGAACCACTAGAACGAGTTGACCAATCAGTACAGCC 255
QY 181 TCTACTTTCATCTCGAATCATGTACCTCTCATCAAGAAAAGATGATGTTGCA 240
DB 256 TCTACTTTCATCTCGAATCATGTACCTCTCATCAAGAAAAGATGATGTTGCA 315
QY 241 CTTGGTGTGATTCAGTCAATCTGCTCCTTCACTCAATGTTTGGCAAGCATG 300
DB 316 CTTGGTGTGATTCAGTCAATCTGCTCCTTCACTCAATGTTTGGCAAGCATG 375
QY 301 ACCCTTCATTCACCAAGAACTGATGATTCACTCACTCTGGGGTTGAGACCCGG 360
DB 376 ACCCTTCATTCACCAAGAACTGATGATTCACTCACTCTGGGGTTGAGACCCGG 435
QY 361 GTCCCTCACTTGGTTCACCAACTCTCTTCTATCTCAATCTGCTCAAGCATATC 420
DB 436 GTCCCTCACTTGGTTCACCAACTCTCTTCTATCTCAATCTGCTCAAGCATATC 495
QY 421 ACCGGATACATG 432
DB 496 ACCGGATACATG 507

RESULT 4

BG524903
 LOCUS BG524903 650 bp mRNA linear EST 16-NOV-2001
 DEFINITION 8-8 Stevia field grown leaf cDNA Stevia rebaudiana cDNA 5', mRNA
 sequence.
 ACCESSION BG524903
 VERSION BG524903.1 GI:16948350
 KEYWORDS EST.
 SOURCE Stevia rebaudiana.
 ORGANISM Stevia rebaudiana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Stevia.
 REFERENCE 1 (bases 1 to 650)
 Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.
 Leaf ESTs from Stevia rebaudiana: A resource for gene discovery in
 diterpene synthesis
 JOURNAL Unpublished (2001)
 COMMENT Contact: Jim Brandle
 Genomics and Biotechnology
 Agriculture and Agri-Food Canada - SCPPRC
 1391 Sandford St., London, Ontario, CANADA, N5V 4T3
 Tel: 519 457 1470
 Fax: 519 457 3997
 Email: brandleje@em.agr.ca
 Seq primer: T3 promoter primer.
 FEATURES
 source
 1..650
 Location/Qualifiers
 /organism="Stevia rebaudiana"
 /strain="751/1501"
 /cultivar="Landrace"
 /db_xref="taxon:35670"
 /clone_lib="Stevia field grown leaf cDNA"
 /tissue_type="leaf"
 /dev_stage="field grown, mid-size"
 /lab_host="E. coli strain XL0R"
 /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI; This
 cDNA library was constructed from polyA+ enriched mRNA
 from field grown leaves. Mid-size actively growing leaves
 were collected and pooled from several plants and frozen
 immediately after harvesting in liquid nitrogen. The cDNA
 was prepared using an XhoI-poly(dT) linker-primer. An
 EcoRI adapter was ligated to the blunt end cDNA and the
 products were digested with EcoRI and XhoI enabling
 directional cloning into the lambda Zap Express vector.
 The library was amplified using the host strain XL1-Blue
 MRF. Mass excision of the library was performed to
 obtain pBK-CMV phagemid clones in the host strain XL0R.
 Single pass DNA sequencing was performed using the T3
 promoter primer: 5' ATTAACTCTCACTAAAGGA 3'. This library
 was constructed by Alex Richman."
 BASE COUNT 171 a 164 c 143 g 171 t
 ORIGIN
 Query Match 25.1%; Score 333.8; DB 12; Length 650;
 Best Local Similarity 69.5%; Freq. No. 4.4e-90;
 Matches 452; Conservative 0; Mismatches 198; Indels 0; Gaps 0;
 Oy 235 TTCACACTTTGGTTGATTCAGATGTCATACCTTTCACATCAATGCTTGGCGGA 294
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1 TTCAGCGATATGTTGAGAGTCGGCGTGTGTTGGCAACCTTACCGAGTGTTCGCCAT 60
 Oy 295 CGCATGACCCCTTCATTACCAAGAACTCGATGATTACTGACACCTCCGTGGTGG 354
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 61 CGTATGACCGCTTACTATGACCGAGATGTCATGATTACAAAATGTCGCCGAGTGG 120
 Oy 335 ACCGCGGTCCCTGACTTTGGTTCCACAACTCTTCTCTATCTAATCTGCTCAAG 414
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 121 ACTAGAGTGTCTCACTTTGGTTGCTTCACTCAATCGCTTCTACTCTGATCCTTTCAAG 180
 Oy 415 CATATCACCGGATGATGACCGCTGGTATGATTCATTACAAAGCTGGACGCTGAT 474
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 181 CATATACCAACTATGACCACTGATGTCATATACCAACACTGGCTACAAAGAT 240

Oy 475 GGTGAGACTCTGTGTTGGAGCCCTTATGACTTATGATGTCGTAGCTGTGAAGTCA 534
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 241 ACCGAGAAATCTTTGCGAGCTCCATATGATTTTCGTTACGGGTTAGCATCCGAGACAT 300
 Oy 535 CCTTCACAGTGGGTTCCAGAGTTCCCTCAAAAGATCTTAAGAATTTGATAGAACAGCAAG 594
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 301 CCTTGTGATGTGGGCTCACTTACCTCCAAAACCTTAACAGCTTAATCGAAAACCGAAC 360
 Oy 595 AATTCCAATAATGGGAAGCAGTACATCTTCCTCCACAGTTTGGAGAGCCCTATTTGTC 654
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 361 AACACAAATGGGCGACTCCCGGTGATCTCTGCTCCACAGTCTTGGAGGCTGTTCGCA 420
 Oy 655 CTACAACTACTAATAATAGAAACCCCTCTTGGCGCAAAAATTCATCAACACTTCATT 714
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 421 CTCCAACTCTCAACCGTAACCCACACCTCATGGCGCAAAACATATTAACATTTATTC 480
 Oy 715 GCTCTTTCAGCTTCATGGGTTGTGCTATAGACGAATGTACACCTTTCATCTGGCAAC 774
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 481 GGGTGTCCGACCATGGGTTGAACGGTTGACGAGATGTTAAGCTTTCCTTGGGAAC 540
 Oy 775 ACTTTGGAGAGTCCCTTACTGAGACCTTATATAGTAGGAGTGAACAAAGAGCTCCGAG 834
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 541 ACTCTGGAGTTCCTCACTGTTGAACCCGTTGCTGCTCAGAAACAGACAGCGAGTCTGAA 600
 Oy 835 ACTAACCTTTGGCTTTGGCTTAACCCAAAATTTTGGTCTCAAAAAC 884
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 601 ACCAAATTATGGCTCATGCCAGCAAGAAACATTTNCTACACAGACAC 650
 RESULT 5
 BE322181 673 bp mRNA linear EST 14-JUL-2000
 LOCUS NF010H12IN1F1101 Insect herbivory Medicago truncatula cDNA clone
 DEFINITION NF010H12IN 5', mRNA sequence.
 ACCESSION BE322181
 VERSION BE322181.1 GI:9195958
 KEYWORDS
 SOURCE
 ORGANISM
 barrel medic.
 Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 REFERENCE 1 (bases 1 to 673)
 Korth, K., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores
 H.R., Inman, J.T., Weller, J.W. and May, G.D.
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula Insect herbivory library
 JOURNAL Unpublished (2000)
 COMMENT Contact: Korth K
 Dept. of Plant Pathology
 University of Arkansas
 217 Plant Science Building, Fayetteville, AR 72701, USA
 Tel: 501 575 5191
 Fax: 501 575 7601
 Email: kthorth@comp.uark.edu
 Medicago Genome Initiative accession: MGI:s:24007
 Insert length: 673 Std Error: 0.00
 Plate: 010 row: H column: 12
 Seq primer: TCACACAGGAACAGCTATGAC.
 FEATURES
 source
 1..673
 Location/Qualifiers
 /organism="Medicago truncatula"
 /db_xref="taxon:3880"
 /clone="NF010H12IN"
 /clone_lib="Insect herbivory"
 /tissue_type="local and systemic leaves"
 /dev_stage="mature"
 /note="Vector: lambda zap; Library was produced from fully
 expanded M. truncatula leaves of plants fed upon by
 Spodoptera exigua (beet armyworm) for 24 hours. Systemic
 (undamaged leaves from injured plants) and wounded leaves

Db	601	GGGAGCGTGAACATGGCGGCGCTTTGGCG	629
RESULT 7			
LOCUS			
B085578			
DEFINITION	B085578	630 bp	mRNA linear EST 14-AUG-2002
FEATURES	OC627805..yg.ab1 QG-ABCD1	lettuce salinas	Lactuca sativa cDNA clone
ACCESSION	OC627805		mRNA sequence.
VERSION	B085578		
KEYWORDS	B085578.1	GI:22241043	
SOURCE	EST.		
ORGANISM	Lactuca sativa.		
	Lactuca sativa		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;		
	Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;		
	Lactuca.		
REFERENCE	1 (bases 1 to 630)		
AUTHORS	Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,		
	Lin,H., van Damme,M., Lavelle,D., Chevallier,P., Ziegler,J., Ellison		
	,P., Kolman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,		
	Church,S., Jackson,L. and Bradford,K.		
	Lettuce and Sunflower ESTs from the Compositae Genome Project		
	http://compgenomics.ucdavis.edu/		
	unpublished (2002)		
JOURNAL	Contact: Alexander kozik [R.W.Michelmore]		
COMMENT	Department of Vegetable Crops, R.W.Michelmore Lab		
	University of California at Davis (UCD)		
	Asmundson Hall, UCD, Davis, CA 95616, USA		
	Tel: 1-(530)-742-1742		
	Fax: 1-(530)-752-9659		
	Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]		
	belongs to contig QG_Ca_cont19973, see http://cgpub.ucdavis.edu/		
	for details.		
	Plate:Q6827	row: B	column: 05.

FEATURES	Location/Qualifiers
source	1. .630

```

/organism="Nactuca salvia"
/cultivar="salinas"
/db_xref="taxon:4236"
/clone="CG827B05"
/clone_1lb="06_ABCDI lettuce salinas"
/1ab_host="E.coli"
/notes="Vector: pBRCNDNA51AB: The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cpdb.ucdavis.edu/
TAG_1lb-06_ABCDI lettuce salinas
TAG_1TISUE-chemical induction
TAG_SEQ-TGTAGCGGG"

```

Query Match	22.2%	Score 296.2;	DB 14;	Length 630;
Best Local Similarity	66.9%	Pred. No. 1.2e-78;		
Matches 421; Conservative	0;	Mismatches 208;	Indels 0;	Gaps 0

[illegible]

Oy	709	TTTCATTCGCTTTACAGCTTCATGGGGTGGTCATATGACGAATGTACACCTTGCATCT	768
Db	181	TTTCATTCGCACTAGCGGCACCATGCGGGTGGCACGGTTGACGAGATGTTGACTTTTGCATCC	240
Oy	769	GGCAACACTTTTGGGAGTGCCCTAGTGACCCCTTATATATGAGAGGATGAACAAAGACG	828
Db	241	GGGAATTCACGTAGAGATCCCTCCCTCGTAAACCCCTTGCTTTATGAAACCAACACGAGAC	300
Oy	829	TTCGAGAGTAACCTTGTGGCTTTGGCTACACCAAAATTTTGGTCCCTGAAAAACCAATA	888
Db	301	TTAGAGAGCAACTTATATGGCTATTGTCCACGCTACAAACAGTATCCCGAGAGAAACCACTT	360
Oy	889	GTGATTAACCTCAATTAGGCCCTATTACGCTCATGACATGGTTGATTTCTTAAAGACATT	948
Db	361	GTGTTACCAAAACCTCGACTTATTCCTCTTGTGATATTTCAAAGATTTTACAGATATTC	420
Oy	949	GGTTTTTCCTGAAGGGGTTTATCTTATATGAAACACGAATTTCTACCTGTTATAGGAACATA	1008
Db	421	GGCTTTCCCTGAAGGGGTGTGACCTTACCGAAACGCGTCTCCCTTGGTGGAGAAAGTTG	480
Oy	1009	AAAGCACCACAAGTAGCTCTTAACTTGATATTATGGAACGGAAGTGGGAACCTTGAACAACA	1068
Db	481	GCTGACACCAAGAAATTCGCTTACATGATATATTACGGAAGTGGGAGTACAGACGCCGGAATCT	540
Oy	1069	TTGTTTATGGGAAGGTGATTTTGATGACGGCCAGAAATATCATATGGGATGGTGAT	1128
Db	541	TTGCAATTACGGGGAAGAAACGTTTATATAGCAGCCGAGATGTGTACAGAGATGAGAT	600
Oy	1129	GGAACGGTAACCTTGGTAGCTTGTGGC	1157
Db	601	GGGACGGTBAAAATGGCAGACTGTTGGC	629

RESULT	8
BG523648	
LOCUS	BG523648
DEFINITION	BG523648
SEQUENCE.	34-12 Stevia field grown leaf cDNA Stevia rebadiana CDNA 5'
ACCESSION	BG523648 . GI:16947068
VERSION	BG523648.1
EST	16-NOV-2001
mRNA	linear

KEYWORDS	EST.
SOURCE	<i>Stevia rebaudiana</i> .
ORGANISM	<i>Stevia rebaudiana</i> Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asterales; euasterids II; Asterales; Asteraceae; Asteroideae; Heliantheae; <i>Stevia</i> . 1 (phases 1 to 587).
REFERENCE	Branlde,J.E., Richman,A., Swanson,A.K. and Chapman,B.P. Leaf ESTs from <i>Stevia rebaudiana</i> : A resource for gene discovery in diterpene synthesis
JOURNAL	unpublished (2001)
COMMENT	Contact: Jim Branle

Genomics and Biotechnology
Agriculture and Agri-Food Canada - SCFPFC
1391 Sandford St., London, Ontario, CANADA, N5V 4T3
Tel: 519 457 1470
Tel: 519 457 3997
Fax: 519 457 3997
Email: brandejeem.agr.ca
Seq primer: T3 promoter primer.
Location/Qualifiers
1..587

```

/organism="Stevia rebaudiana"
/strain="751/1501"
/cultivar="landrace"
/db_xref="taxon:55670"
/clone_lib="Stevia field grown leaf cDNA"
/tissue_type="leaf"
/dev_stage="field grown, mid-size"
/lab_host="E. coli strain XJ018"
/notes="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI. This
cDNA library was constructed from polyA+ enriched mRNA

```


Db	Accession	Definition	Source	Organism	Reference Authors Title	Journal Comment	Features Source
Db	462	TACAAAGATACGAAACCTTTTCGAGATCCATATGATTTTCGTTACGGTTAGATCC	521				
0y	526	GAAGGTCAACCTTCACAACTGGGTTCCAAAGTTCCCTCAAAAGATCTAAAGCAATTTGATAGAA	585				
Db	522	GAGGACCACTTGTGATGTGGCTCAACTTACCTCCAAACCAAGCTTAAGCAATTAAACGAA	581				
0y	586	GAGCAACCA 596					
Db	582	AACGCAACCA 592					
RESULT 10	BG522086	635 bp	mRNA	linear	EST 01-FEB-2002		
LOCUS	BG522086						
DEFINITION	18-18 Stevia field grown leaf cDNA	Stevia rebaudiana cDNA 5', mRNA					
ACCESSION	BG522086						
VERSION	BG522086.1	GI:18465150					
KEYWORDS	EST.						
SOURCE	Stevia rebaudiana.						
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae; Heliantheae; Stevia.						
REFERENCE	1 (bases 1 to 635)						
AUTHORS	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.						
TITLE	Leaf ESTs from Stevia rebaudiana: A resource for gene discovery in diene synthase						
JOURNAL	Unpublished (2001)						
COMMENT	Contact: Jim Brandle Genomics and Biotechnology Agriculture and Agri-Food Canada - SCPPFC 1391 Sandford St., London, Ontario, CANADA, N5V 4T3 Tel: 519 457 1470 Fax: 519 457 3997 Email: brandle.j@em.agr.ca Seq primer: T3 promoter primer. Location/Qualifiers						
FEATURES	1. 635						
Source	/organism="Stevia rebaudiana"						
	/strain="751/1501"						
	/cultivar="landrace"						
	/db_xref="taxon:55670"						
	/clone_lib="Stevia field grown leaf cDNA"						
	/tissue_type="leaf"						
	/dev_stage="field grown, mid-size"						
	/lab_host="E. coli strain XL0R"						
	/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from polyA+ enriched mRNA from field grown leaves. Mid-size actively growing leaves were collected and pooled from several plants and frozen immediately after harvesting in liquid nitrogen. The cDNA was prepared using an XhoI-poly(dT) linker-primer. An EcoRI adapter was ligated to the blunt end cDNA and the products were digested with EcoRI and XhoI enabling directional cloning into the lambda ZAP Express vector. The library was amplified using the host strain XL1-Blue MRF. Mass excision of the library was performed to obtain pBK-CMV phagemid clones in the host strain XL0R. Single pass DNA sequencing was performed using the T3 promoter primer: 5' ATTAACTCTCTAAGGCA 3'. This library was constructed by Alex Richman."						
BASE COUNT	158 a	157 c	164 g	155 t	1 others		
ORIGIN							
Query Match	17.6%	Score 234.6;	DB 12;	Length 635;			
Best Local Similarity	67.3%	Pred. No. 7.1e-60;					
Matches 330;	Conservative 0;	Mismatches 160;	Indels 0;	Gaps 0;			
0y	106	CTCGACCTTTGATCTCTAATACAGGTAAAGGAGCAACCAACTGAAAGCAAGTTGACG	165				
Db	115	CTGATTCGGTGATCTCTGCTGACCGCGTGGCAACCAAGTAAAGCCCGTTAACG	174				

Accession	Source	Organism	Reference Authors Title	Journal Comment
166	AAAGATGACAGCCCTACTTTCATCTGCGAATCATGTATGCCCTTCATCAAGAAAG	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
175	GCGGAGTATTAAGGGGACTACCTGCTGTCTCAACCGCTTCTACCCGCTCAAGAGATGAG	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
226	AATGATGGTTGACAGCTTTGGTTTGAATTCACAGTGCATATCTTCTCTCACTCAATGC	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
235	GAGGAGGTGTTGACAGCTATGGTTGCGAGGTGCGCGTGTGGGACCAATGACCGAGTGT	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
286	TTTGGCCGACAGCATGACCTTCATTTACCAACCAAGAACTGCATGATTTACTTCAACACTCT	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
295	TTTGTCCATGTTATGACGCTTTACTATGACCGAAGATGTCGATGATTTACAAAATGCTCC	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
346	GAGGTTGAGACCGGGTCCCTCACTTTGGTTTGCACCACTCTCTCTATCTCAATCT	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
355	GAGATGAGACTAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
406	CGTCTCAAGCATATACCGGATACATGACGACCCCTGTGATGATTTACAAACCTTGGC	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
415	TCCTTCAAGCATTTAACAACATACATGACGACCATTTGGGATCATTTACCAACACTTGGC	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
466	TACGCTGATGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
475	TACAAATATCCGAGATCTTTGAGAGTCCATATGATTTTCGTTACGATGAGGTTTACATCC	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
526	GAAGGTCACCTTCACAGTGGGTTTCCAAATTCCTCAAGATCTTAAGAAATTTGATGAA	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
535	GAGGAGATCTCTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
586	GAGGAGATCTCTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
595	AAGCAGCAAGA 595	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle
604	AAGCAGCAAGA 604	Stevia rebaudiana	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.	Contact: Jim Brandle

from field grown leaves. Mid-size actively growing leaves were collected and pooled from several plants and frozen immediately after harvesting in liquid nitrogen. The cDNA was prepared using an XhoI-poly(dT) linker-primer. An EcoRI adapter was ligated to the blunt end cDNA and the products were digested with EcoRI and XhoI enabling directional cloning into the lambda ZAP Express vector. The library was amplified using the host strain XL1-Blue MR⁺. Mass excision of the library was performed to obtain pBR-CMV phagemid clones in the host strain XL0R. Single pass DNA sequencing was performed using the T3 promoter primer: 5' ATTAACTCTCACTAAAGGCA 3'. This library was constructed by Alex Richman."

BASE COUNT 168 a 117 c 162 g 161 t 8 others
ORIGIN

Query Match 16.3%; Score 216.8; DB 12; Length 616;
Best Local Similarity 64.4%; Pred. No. 1.8e-54;

Matches 337; Conservative 0; Mismatches 185; Indels 1; Gaps 1;

QY 671 GAATCCCCCTCTGGCGGCAAAATTCATCAACACTTATGCTTTCAGCTCAT 730
DB 4 GTAAACCCCTCTGCGGCAAAATTCATCAACACTTATGCTTTCAGCTCAT 63
QY 731 GGGGTGGTGTATAGAGCAATGTACACCTTGCATCTGCACACCTTGGG-AGTGCC 789
DB 64 GGGGTGGAGCGTTGAGAGATTTTAACTTTGCTTGGGACACCTTGGGAGTTCA 123
QY 790 CTAGTGAACCTTTATTAGTGGAGATGAACAAGCTCCGAGAGTACCTTTGGCTT 849
DB 124 CTGTTGAACCCGTTGCTGTCAGAAATGACACGAGTTCTGAAGCATTTATGGCC 183
QY 850 TTGGCTTAACCAAAATTTTGGCTCAAAACCAATAGTATAGTCAATAGGCTT 909
DB 184 ATGCCACGACGAAACATTTTCTCAACAGACCTGTAGTTGTCATGATGATCACT 243
QY 910 TATTCAGCTCATGATGATGTTTCTTAAAGACATTTGTTTCCGTAAGGGGTTTAT 969
DB 244 TACCTGCTTTTTCATCTTACAGGTTCTTAAATATCGGTTTCAAGAGGCTGTGAC 303
QY 970 CTTTATGAACAGCAATTTTCACTTGTATAGGAACATTAAGACCAACAACCTTAT 1029
DB 304 CTTTACGACAGACGATCTTGTGCTTGTGTAAGATTTGTTGACCGGAGTGCCGTA 363
QY 1030 ACTTGTATTATGGAACGGAGTGGGAACCTTGAACATTTTATGGAAGGTGAT 1089
DB 364 ACGGTATAGTGGAAACGGTGTCCGACCGGAACTTTGTTATTTGGGAGACAGCGG 423
QY 1090 TTTGATGAACGGCCAGAAATATCATATGGGATGTTGATGACGAGTGTGAGC 1149
DB 424 TTGATATAGCAGCGGAAATTTGTTATGAGATGTTGACGAGACGTTAATATGTTAGG 483
QY 1150 TTGTTGGCCTTCAATCATATGGAAGAGGAGAAATCAAT 1192
DB 484 TTGTTGGCTTTGGAGAGCAAGTGAAGGATGAATAAATCAAT 526

RESULT 12
B0864610 527 bp mRNA linear EST 14-AUG-2002
LOCUS OGC27C22.yg.abl OG_ABCDI lettuce salinas Lactuca sativa cDNA clone
DEFINITION OGC27C22. mRNA sequence.
ACCESSION B0864610
VERSION B0864610.1 GI:22250075
KEYWORDS EST.
SOURCE Lactuca sativa.
ORGANISM Lactuca sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; eunasterids II; Asterales; Asteraceae; Lactuceae;
Lactuca.

REFERENCE 1 (bases 1 to 527)
AUTHORS Kozik,A., Michlemore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,

TITLE
JOURNAL
COMMENT
Contact: Alexander Kozik [R.W. Michlemore]
Department of Vegetable Crops, R.W. Michlemore Lab
University of California at Davis (UCD)
Asmudson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.edu [michlemore@ucdavis.edu]
singletons, see http://cgpbdb.ucdavis.edu/ for details.
Plate: OGC27 row: C column: 22.

FEATURES

source

Location/Qualifiers
1. 527
/organism="Lactuca sativa"
/cultivar="Salinas"
/db_xref="taxon:4236"
/clone="OGC27C22"
/clone_id="OG_ABCDI lettuce salinas"
/lab_host="E.coli"
/note="Vector: pBRCDNA51AB: The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpbdb.ucdavis.edu/TAG-LIB-OG_ABCDI lettuce salinas
TAG-TISSUE-flowers post-fertilized
TAG_SEQ-TGCCATCGCG"

BASE COUNT 122 a 125 c 147 g 133 t
ORIGIN

Query Match 15.4%; Score 205.2; DB 14; Length 527;
Best Local Similarity 68.2%; Pred. No. 5.7e-51;

Matches 285; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 104 ACCGACACCTTTGATTTATACAGTAAGGAGGAGACCACTAGACAGGTTGA 163
DB 110 AGCTTATCCGTTACTTGTGCTCCGAGCGGCGGAGACCACTAGACAGGTTGA 169
QY 164 CCAATCACTACAGCCCTCTACTTTCATCTCGAATCATGTGTACCTCTCAAGAAA 223
DB 170 CGGGGAGTATTAAGGAGAGCGAGTGTGATGAGCCGATTTACCAATTAAGAGCAG 229
QY 224 AGATGATGTTGACACTTTGTTGATTCAGAGTGTATCTGCTCTTCACTCAT 283
DB 230 AGGGAGGTGTTGCTCCCTATGCTGAGCGTGTATTTATGACCACTTCTTGGAGT 289
QY 284 GCTTGGCCGAGAGATGACCTTCATTTACCAACAAGATGATGATTTACTCAACACT 343
DB 290 GTTTTGGTGAAGATGAGCGCTTTACTACCAACGAATTTGATGATTTACAGAAATGCTC 349
QY 344 CTGGGTTGAGACCGGCTCCCTCACTTTGTTCCACCACTCTCTTCTATCTCAATC 403
DB 350 CGGGAGTTGAGACGAGGCTGTGTAGTTGTTTCCACATCACTCTTCTACCTGATC 409
QY 404 CTGCTGTCACATATACCGGATATGACACCCCTGTGATTTACATTAAGAACTTG 463
DB 410 CTAATCTCAACACATATACGATTAACATGAACCACTTGTGAGTCTCTCAAGAGCTTG 469
QY 464 GCTACGCTGATGAGAGCTGCTGTTGGAGCCCTTATGATTTAGATGATGCTGAC 521
DB 470 GCTACACGATTAACAAACATTTGCGAGACCTTACGATTTCCGTTATGCGCTTGC 527

RESULT 13
AV441327/c

[illegible]

QY	1159	CTTCATCATCTATGGAAAGAGAGAAATAATCATCTAACGTTAAAGCTGTTAGATAGATGGG	1218
Db	138	TT-----GAAAGTCGATAGCTTTGGAACACCGCTAGCATTTGATGGA	100
QY	1219	GTGTCTCATCTTCTTCATACCTTAAGAGATGAAGTTGCATTAATGAATAGTACGTGAGATT	1278
Db	99	GTTTCGATACATCTATCTTAAAGACGAGATCGACATTAAGAGATTTATGAAGCAGATT	40
QY	1279	ACTTCATTTAATTTCTCA	1295
Db	39	TCAATTATTAAATTATGGA	23
RESULT 14			
AV827187			
LOCUS	AV827187	RAFL9 Arabidopsis thaliana cDNA clone RAFL09-12-E12 5',	
DEFINITION	AV827187	mRNA sequence.	
ACCESSION	AV827187		
VERSION	AV827187.1	GI:19869247	
KEYWORDS			
SOURCE			
ORGANISM		thale cress.	
REFERENCE		Arabidopsis thaliana	
AUTHORS		Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta; Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae; eurosids II: Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 600)	
		Seki, M., Narusaka, M., Ishida, J., Kamuya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinozaki, K., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.	
TITLE		Large scale analysis of Arabidopsis full-length cDNA (2002b)	
JOURNAL		Unpublished (2002)	
COMMENT		Contact: Motoaki Seki Plant Functional Genomics Research Group RIKEN Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-4359 Fax: 81-298-36-9060 Email: mseki@rtc.riken.go.jp An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified lambda FliC1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/Plant/index_e.html) for further details.	
FEATURES			
source		Location/Qualifiers	
		1..600	
		/organism="Arabidopsis thaliana"	
		/db_xref="taxon:3702"	
		/clone="RAFL09-12-E12"	
		/clone_id="RAFL9"	
		/dev_stage="plants at various developmental stages from germination to mature seeds"	
		/lab_host="DH10B"	
		/note="Site_1: BamHI. Site_2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"	
BASE COUNT	158 a	129 c	136 g 174 t 3 others
ORIGIN			
Query Match	12.9%	Score 171.6;	DB 10; Length 600;
Best Local Similarity	62.7%	Pos. No. 1..40;	
Matches 299; Conservative	0;	Mismatches 172;	Indels 6; Gaps 2;
QY	45	ACGTACAGTTACAGAGTGTGTGATCTCTGCATGTCGTAGCAATGTGGGCAAGCA	104
Db	124	AGTCAATACGATCTCGTTGTGTGATCATGCTGATGATCAAGCTGTGTGGATGCA	183
QY	105	CCGTGACCTTTGATTTCAATACAGGTAAGGAGGAACCACTAGAGCAAGTTGAC	164

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 19:11:35 ; Search time 40 Seconds
(without alignments)
1475.751 Million cell updates/sec

Title: us-09-857-612a-14

Perfect score: 2333
Sequence: 1 MKKEQERGKLEIVATLVTV.....GETTSINSHALGSLNFG 443

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: A_Geneseq-101002.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2333	100.0	443	21	soybean putative 1
2	2022	86.7	381	21	soybean putative 1
3	1380.5	59.2	432	21	Arabidopsis thaliana
4	1380.5	59.2	432	21	Arabidopsis thaliana
5	1380.5	59.2	432	21	Arabidopsis thaliana
6	1380.5	59.2	432	21	Arabidopsis thaliana
7	1380.5	59.2	432	21	Arabidopsis thaliana
8	1367	58.6	413	21	Arabidopsis thaliana
9	1367	58.6	413	21	Arabidopsis thaliana
10	1366	58.6	410	21	Arabidopsis thaliana

11	1366	59.6	410	21	AA642648	Arabidopsis thaliana
12	1345	57.7	387	21	AA624263	Arabidopsis thaliana
13	913.5	39.2	417	21	AA601209	Corn putative leci
14	739.5	31.7	439	21	AA601210	Corn putative leci
15	717	30.7	434	21	AA601208	Corn putative leci
16	612	26.2	412	19	AA60944	Murine kidney leci
17	610	26.1	412	19	AA60942	Human heart leci
18	610	26.1	412	20	AA641708	Human heart leci
19	610	26.1	412	21	AA644264	Human PRO540 prote
20	610	26.1	412	21	AA619579	Human PRO540 used
21	610	26.1	412	22	AA693359	Human polypeptide,
22	610	26.1	412	22	AA693243	Human protein sequ
23	610	26.1	412	22	AA650954	Human PRO540 prote
24	610	26.1	412	23	AB634036	Human Pro peptide
25	608	26.1	379	19	AA680937	Human heart leci
26	608	26.1	382	19	AA680940	Human heart leci
27	606	26.0	379	19	AA680939	Murine leci
28	596	25.5	444	19	AA680943	Human kidney leci
29	594	25.5	411	19	AA680938	Human kidney leci
30	594	25.5	414	19	AA680941	Human kidney leci
31	560	24.0	421	19	AA667428	Drosophila melanog
32	543	23.3	440	8	AA670134	Natural recombinan
33	541	23.2	440	18	AA624789	Human leci
34	541	23.2	440	20	AA649562	Human leci
35	541	23.2	440	22	AA672627	Human leci
36	541	23.2	440	23	AA699745	Human leci
37	541	23.2	440	23	AA684290	Human leci
38	541	23.2	440	23	AA651918	Human leci
39	518.5	22.2	318	21	AA691645	Anti-ICAM antibody
40	493	21.1	187	21	AA618062	P. radiata leci
41	421.5	18.1	233	21	AA601206	Corn putative leci
42	376	16.1	117	21	AA618125	P. radiata leci
43	368.5	15.8	143	21	AA601205	Corn putative leci
44	225	9.6	281	22	AB627613	Novel human diagno
45	222.5	9.5	447	22	AA600461	Arabidopsis thaliana

ALIGNMENTS

RESULT 1
ID AAB01211 standard; Protein: 443 AA.
AC AAB01211;
DT 12-DEC-2000 (first entry)
DE Soybean putative leci:cholesterol acyltransferase #2.
KW Soybean; lecithin:cholesterol acyltransferase; phytosterol;
KW phosphatidylcholine-sterol O-acyltransferase; heat shock; cold shock.
OS Glycine max.
PN W0200032791-A2.
PD 08-JUN-2000.
PF 02-DEC-1999; 99MO-US28586.
PK 03-DEC-1998; 98US-0110782.
XX (DUPO) DU PONT DE NEMOURS & CO E. I.
XX Cahoon RE, Kinney AJ, Sakai H, Shen JB, Butler KH, Saylor JT;
XX WPT; 2000-412337/35.
XX N-PSDB; AAA49206.
XX Polynucleotide encoding plant lecithin:cholesterol acyltransferase
XX enzyme useful for producing transgenic plants and for producing
XX antibodies specific to which is useful for screening cDNA expression

PT Libraries -
XX
PS Claim 10: Page 46-47; 49pp; English.
XX
CC The present sequence is a putative protein sequence of a soybean
CC lecithin:cholesterol acyltransferase (also known as
CC phosphatidylcholine-sterol O-acyltransferase). This enzyme is found
CC associated with high-density lipoproteins and Apolipoprotein-AI and -D.
CC The gene and protein can be used to produce transgenic plants which have
CC increased lipid metabolism and membrane fluidity, and therefore increased
CC resistance to heat and/or cold shock, to alter the content of phytosterol
CC or lecithin in grains and to identify potential herbicides.
XX
SQ Sequence 443 AA;
Query Match 100.0%; Score 2333; DB 21; Length 443;
Best Local Similarity 100.0%; Pred. No. 4.1e-219;
Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKKQEGELKIEVATLVTVVVMVMSLCTCGASNDLPILIPGNGNQLBARLTNOYKP 60
DQ 1 MKKQEGELKIEVATLVTVVVMVMSLCTCGASNDLPILIPGNGNQLBARLTNOYKP 60
QY 61 STFCESWYPLIKKNGWFLMFDSSVILAFPTQCFARMTLHYHQELDDYFNTPGVETR 120
DQ 61 STFCESWYPLIKKNGWFLMFDSSVILAFPTQCFARMTLHYHQELDDYFNTPGVETR 120
QY 121 VPHGGSNLSILYNPRLKHTGYMAPLVDSLOKLGADGETLFGAPYDFRYGLAEGHPS 180
DQ 121 VPHGGSNLSILYNPRLKHTGYMAPLVDSLOKLGADGETLFGAPYDFRYGLAEGHPS 180
QY 181 QVGSKFLDKNLIEASNSNNGKPVILSHSLGFLVQLNRPSPMRKKFIKHFIAL 240
DQ 181 QVGSKFLDKNLIEASNSNNGKPVILSHSLGFLVQLNRPSPMRKKFIKHFIAL 240
QY 241 SAPMGAIDEMWTFAAGNTLGVPLVDPLVVRDEQRSSSNMLLPPKIFGPQPIVITP 300
DQ 241 SAPMGAIDEMWTFAAGNTLGVPLVDPLVVRDEQRSSSNMLLPPKIFGPQPIVITP 300
QY 301 IRPASHDMVDFLKDIGFPEGVYPERIRILPLIGNIKAPQVPTICIMGTGVTLETFYG 360
DQ 301 IRPASHDMVDFLKDIGFPEGVYPERIRILPLIGNIKAPQVPTICIMGTGVTLETFYG 360
QY 361 KGDERPERISYGDGDTVNLVSLALQSLMKKEKNQYLKVVKIDGVSHTSILKDEVALN 420
DQ 361 KGDERPERISYGDGDTVNLVSLALQSLMKKEKNQYLKVVKIDGVSHTSILKDEVALN 420
QY 421 EIVGEITSINSHAEGLSNLFSG 443
DQ 421 EIVGEITSINSHAEGLSNLFSG 443
RESULT 2
AAB01207
ID AAB01207 standard; Protein: 381 AA.
XX
AC AAB01207;
XX
DT 12-DEC-2000 (first entry)
XX
DE Soybean putative lecithin:cholesterol acyltransferase #1.
XX
KM Soybean: lecithin:cholesterol acyltransferase; phytosterol;
KM phosphatidylcholine-sterol O-acyltransferase; heat shock; cold shock.
XX
OS Glycine max.
XX
PN WO200032791-A2.
XX
PD 08-JUN-2000.
XX
PF 02-DEC-1999; 99MO-US28586.
XX

PR 03-DEC-1998; 98US-0110782.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E. I.
XX
PI Cahoon RE, Kinney AJ, Sakai H, Shen JB, Butler KH, Saylor JJ;
XX
DR WPI: 2000-412337/35.
XX
DR N-PSDB: AAA49202.
XX
PT Polynucleotide encoding plant lecithin:cholesterol acyltransferase
PT enzyme useful for producing transgenic plants and for producing
PT antibodies specific to which is useful for screening cDNA expression
PT libraries -
PS Claim 10: Page 38-39; 49pp; English.
XX
XX The present sequence is a putative protein sequence of a soybean
CC lecithin:cholesterol acyltransferase (also known as
CC phosphatidylcholine-sterol O-acyltransferase). This enzyme is found
CC associated with high-density lipoproteins and Apolipoprotein-AI and -D.
CC The gene and protein can be used to produce transgenic plants which have
CC increased lipid metabolism and membrane fluidity, and therefore increased
CC resistance to heat and/or cold shock, to alter the content of phytosterol
CC or lecithin in grains and to identify potential herbicides.
XX
SQ Sequence 381 AA;
Query Match 86.7%; Score 2022; DB 21; Length 381;
Best Local Similarity 100.0%; Pred. No. 8.4e-189;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 63 FICSWYPLIKKNGWFLMFDSSVILAFPTQCFARMTLHYHQELDDYFNTPGVETR 122
DQ 1 FICSWYPLIKKNGWFLMFDSSVILAFPTQCFARMTLHYHQELDDYFNTPGVETR 122
QY 123 HFGSTNSILYNPRLKHTGYMAPLVDSLOKLGADGETLFGAPYDFRYGLAEGHPSQV 182
DQ 123 HFGSTNSILYNPRLKHTGYMAPLVDSLOKLGADGETLFGAPYDFRYGLAEGHPSQV 182
QY 183 GSKFLDKNLIEASNSNNGKPVILSHSLGFLVQLNRPSPMRKKFIKHFIALSA 242
DQ 183 GSKFLDKNLIEASNSNNGKPVILSHSLGFLVQLNRPSPMRKKFIKHFIALSA 242
QY 243 PMGGAIDEMWTFAAGNTLGVPLVDPLVVRDEQRSSSNMLLPPKIFGPQPIVITP 302
DQ 243 PMGGAIDEMWTFAAGNTLGVPLVDPLVVRDEQRSSSNMLLPPKIFGPQPIVITP 302
QY 303 PYSADHMDVDFLKDIGFPEGVYPERIRILPLIGNIKAPQVPTICIMGTGVTLETFYGK 362
DQ 303 PYSADHMDVDFLKDIGFPEGVYPERIRILPLIGNIKAPQVPTICIMGTGVTLETFYGK 362
QY 363 DEDERPERISYGDGDTVNLVSLALQSLMKKEKNQYLKVVKIDGVSHTSILKDEVALNEI 422
DQ 363 DEDERPERISYGDGDTVNLVSLALQSLMKKEKNQYLKVVKIDGVSHTSILKDEVALNEI 422
QY 423 VGEITSINSHAEGLSNLFSG 443
DQ 423 VGEITSINSHAEGLSNLFSG 443
RESULT 3
AAB24260
ID AAB24260 standard; Protein: 432 AA.
XX
AC AAB24260;
XX
DT 12-FEB-2001 (first entry)
XX
DE Arabidopsis thaliana PDAT amino acid sequence SEQ ID NO:14.
XX
KM PDAT: phospholipid:diacylglycerol acyltransferase; triacylglycerol;
KM TAG; EST: expressed sequence tag; fatty acid; oil content.
XX

OS Arabidopsis thaliana.
 XX
 PN WO20060095-A2.
 PD 12-OCT-2000.
 XX
 XX
 PE 28-MAR-2000; 2000WO-EP02701.
 XX
 XX 01-APR-1999; 99EP-0106656.
 PR 10-JUN-1999; 99EP-0111321.
 PR 07-FEB-2000; 2000US-0180687.
 XX
 PA (BADI) BASF PLANT SCI GMBH.
 XX
 PI Dahlglyst A, Stahl U, Lenman M, Banas A, Ronne H, Steyme S;
 XX
 XX WPI: 2000-665012/64.
 XX
 XX Phospholipid:diacylglycerol acyltransferase enzymes in the biosynthetic
 PT pathway for triacylglycerol production and DNA encoding them, useful
 PT for producing triacylglycerol, or for transforming any cell or organism
 PT to increase oil content -
 XX
 PS Claim 5; Page 64; 97pp; English.
 XX
 XX The present invention describes an enzyme for catalysing (in an
 CC acyl-CoA-independent reaction) the transfer of fatty acids from
 CC phospholipids to diacylglycerol in the biosynthetic pathway for the
 CC production of triacylglycerol (TAG). The enzyme is designated as
 CC phospholipid:diacylglycerol acyltransferase (PDAT). The enzyme and
 CC the nucleotides encoding them are useful for producing TAG and/or
 CC TAG with uncommon fatty acids. The enzyme and the nucleotide are
 CC also useful for transforming any cell or organism in order to be
 CC expressed in this cell or organism and result in an altered, preferably
 CC increased oil content of this cell or organism. The present sequence
 CC represents an Arabidopsis thaliana PDAT amino acid sequence.
 XX
 SQ Sequence 432 AA;
 Query Match 59.2%; Score 1380.5; DB 21; Length 432;
 Best Local Similarity 61.0%; Pred. No. 5.3e-126;
 Matches 264; Conservative 62; Mismatches 92; Indels 15; Gaps 5

Db	Accession	Score	Length	Gap
Db	409 ALKMTKOISITN 421			
Db	AA01212 standard; Protein: 432 AA.			
Db	AA01212; 12-DEC-2000 (first entry)			
Db	Arabidopsis thaliana lecithin:cholesterol acyltransferase #1.			
Db	lecithin:cholesterol acyltransferase; phytosterol; phosphatidylcholine-sterol O-acyltransferase; heat shock; cold shock.			
Db	Arabidopsis thaliana.			
Db	WO20032791-A2.			
Db	08-JUN-2000.			
Db	02-DEC-1999; 99WO-US28586.			
Db	03-DEC-1998; 98US-0110782.			
Db	(DUPO) DU PONT DE NEMOURS & CO E. I.			
Db	Cahoon RE, Kinney AJ, Sakai H, Shen JB, Butler KH, Saylor JJ; WPI: 2000-412337/35.			
Db	Polynucleotide encoding plant lecithin:cholesterol acyltransferase enzyme useful for producing transgenic plants and for producing antibodies specific to which is useful for screening cDNA expression libraries -			
Db	Disclosure: Page 47-49; 49pp; English.			
Db	The present sequence is the protein sequence of Arabidopsis thaliana lecithin:cholesterol acyltransferase (also known as phosphatidylcholine-sterol O-acyltransferase). This enzyme is found associated with high-density lipoproteins and Apolipoprotein-AI and -D. The gene and protein can be used to produce transgenic plants which have increased lipid metabolism and membrane fluidity, and therefore increased resistance to heat and/or cold shock, to alter the content of phytosterol or lecithin in grains and to identify potential herbicides.			
Db	Sequence 432 AA;			
Db	Query Match 59.2%; Score 1380.5; DB 21; Length 432; Best Local Similarity 61.0%; Pred. No. 5.3e-126; Mismatches 92; Indels 15; Gaps 5.			
Db	Matches 264; Conservative 62; Mismatches 92; Indels 15; Gaps 5.			
Db	1 MKKEDEGLKIEVATLFTVYVVVVMISLCTGCASNDPLILIPGNGNLEARLTNQYP 60			
Db	1 MKK-----ISSHYSVVIALVYVYVMTSMQANGSNVPIILIPGNGNLEVRIDREYKP 55			
Db	61 STFICESW-YPLIKKNGNGFRLMDESSVYLAFTQCFERMTLHYHOELDDYFNTPGVET 119			
Db	56 SSWCWSMILYPIHKKSGGFRMLPDAAYLLSPFCFSDRMMLYYDDPDDYONAPGYQT 115			
Db	120 RYPHHGSGNSLILYLPRLKHTGYAFLVDSIQ-KLGVADGETLGGAPIDFRRYGLAAE 178			
Db	116 RYPHHGSGNSLILYLPRLDRTATSYMEHLVKALEKKGCVYNDOTILGAPYDFRYGLAAS 175			
Db	179 PSQVSKFLKLDKNTLIEEASNSNNKSPVYLLSHSGGLGVFLDILNRNPSSKKEFLKAPT 238			
Db	176 PSRVASQFLDQLKQVKEKTSSENEKRPVYLLSHSGGLGVFLHFLNRTTPSWAKKIKIHFV 235			
Db	239 ALSAPWGALDEMYTFASGNTLGVPVLVDPFLVDRDORSESSNMULLPNKIFGPQ-KPIV 297			
Db	236 ALAAPWGGLTISOMKTFASGNTLGVPVLVDPFLVDRDORSESSNMULLPNKIFGPQ-KPIV 295			

QY	298	ITPRPISAHMVDPLKDIGCPPEGVYPERITLPLIGNIKAPQVPTICMGTCVILETL	357
Db	296	VTPLPVNTATAYEMDMFFADIFSGQGVVPYKTRVLPLEETLMPGVPPVCIYGRGDPEVL	355
QY	358	FYGGGDFDEREPEIYXGDDGTVNLVSLALQSLKKEKNQYLKVVKIDGVSHSILKDEV	417
Db	356	MYGGGDFPKQPEIKYKQDDGTVNLASLAL-----KVDSLNTVEIDGVSHSILKDEI	408
QY	418	ALNEIVEGETSIN	430
Db	409	ALKEIMKOISIIN	421
RESULT 5			
AAG22979			
ID	AAG22979 standard; Protein; 432 AA.		
XX			
AC	AAG22979;		
XX			
DT	17-OCT-2000 (first entry)		
XX			
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 26114.		
XX			
KW	Protein identification; signal transduction pathway; metabolic pathway;		
KM	hybridisation assay; genetic mapping; gene expression control; promoter;		
XX	termination sequence.		
OS	Arabidopsis thaliana.		
XX			
PN	EP1031405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-0301439.		
XX			
PR	25-FEB-1999;	9905-0121825.	
PR	05-MAR-1999;	9905-0123180.	
PR	09-MAR-1999;	9905-0123548.	
PR	23-MAR-1999;	9905-0125788.	
PR	25-MAR-1999;	9905-0126264.	
PR	29-MAR-1999;	9905-0126785.	
PR	01-APR-1999;	9905-0127462.	
PR	06-APR-1999;	9905-0128234.	
PR	08-APR-1999;	9905-0128714.	
PR	16-APR-1999;	9905-0129845.	
PR	19-APR-1999;	9905-0130077.	
PR	21-APR-1999;	9905-0130449.	
PR	23-APR-1999;	9905-0130510.	
PR	23-APR-1999;	9905-0130891.	
PR	28-APR-1999;	9905-0131449.	
PR	30-APR-1999;	9905-0132048.	
PR	30-APR-1999;	9905-0132407.	
PR	04-MAY-1999;	9905-0132484.	
PR	05-MAY-1999;	9905-0132485.	
PR	06-MAY-1999;	9905-0132486.	
PR	07-MAY-1999;	9905-0132487.	
PR	11-MAY-1999;	9905-0132863.	
PR	11-MAY-1999;	9905-0134256.	
PR	14-MAY-1999;	9905-0134218.	
PR	14-MAY-1999;	9905-0134219.	
PR	14-MAY-1999;	9905-0134321.	
PR	14-MAY-1999;	9905-0134370.	
PR	18-MAY-1999;	9905-0134376.	
PR	19-MAY-1999;	9905-0134471.	
PR	20-MAY-1999;	9905-0135124.	
PR	21-MAY-1999;	9905-0135353.	
PR	24-MAY-1999;	9905-0135629.	
PR	25-MAY-1999;	9905-0136021.	
PR	27-MAY-1999;	9905-0136392.	
PR	28-MAY-1999;	9905-0136782.	
PR	01-JUN-1999;	9905-0137222.	
PR	03-JUN-1999;	9905-0137528.	

PR	04-JUN-1999	9905-0137502
PR	07-JUN-1999	9905-0137722
PR	08-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138540
PR	10-JUN-1999	9905-0138847
PR	14-JUN-1999	9905-0139151
PR	16-JUN-1999	9905-0139452
PR	16-JUN-1999	9905-0139453
PR	17-JUN-1999	9905-0139492
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139456
PR	18-JUN-1999	9905-0139457
PR	18-JUN-1999	9905-0139458
PR	18-JUN-1999	9905-0139459
PR	18-JUN-1999	9905-0139460
PR	18-JUN-1999	9905-0139461
PR	23-JUN-1999	9905-0140593
PR	23-JUN-1999	9905-0140594
PR	24-JUN-1999	9905-0140695
PR	28-JUN-1999	9905-0140823
PR	29-JUN-1999	9905-0140931
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0142877
PR	01-JUL-1999	9905-0142154
PR	02-JUL-1999	9905-0142055
PR	06-JUL-1999	9905-0142390
PR	08-JUL-1999	9905-0142803
PR	09-JUL-1999	9905-0142820
PR	16-JUL-1999	9905-0144086
PR	19-JUL-1999	9905-0144325
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144334
PR	19-JUL-1999	9905-0144335
PR	20-JUL-1999	9905-0144352
PR	20-JUL-1999	9905-0144632
PR	20-JUL-1999	9905-0144684
PR	21-JUL-1999	9905-0144614
PR	21-JUL-1999	9905-0145016
PR	21-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145192
PR	22-JUL-1999	9905-0145143
PR	23-JUL-1999	9905-0145218
PR	23-JUL-1999	9905-0145524
PR	26-JUL-1999	9905-0145726
PR	27-JUL-1999	9905-0145513
PR	27-JUL-1999	9905-0145519
PR	28-JUL-1999	9905-0145551
PR	02-AUG-1999	9905-0146386
PR	02-AUG-1999	9905-0146388
PR	03-AUG-1999	9905-0146389
PR	04-AUG-1999	9905-0147038
PR	04-AUG-1999	9905-0147302
PR	05-AUG-1999	9905-0147304
PR	05-AUG-1999	9905-0147760
PR	05-AUG-1999	9905-0147920

PR	04-JUN-1999	9905-0137502
PR	07-JUN-1999	9905-0137722
PR	08-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138540
PR	10-JUN-1999	9905-0138847
PR	14-JUN-1999	9905-0139151
PR	16-JUN-1999	9905-0139452
PR	16-JUN-1999	9905-0139453
PR	17-JUN-1999	9905-0139492
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139456
PR	18-JUN-1999	9905-0139457
PR	18-JUN-1999	9905-0139458
PR	18-JUN-1999	9905-0139459
PR	18-JUN-1999	9905-0139460
PR	18-JUN-1999	9905-0139461
PR	23-JUN-1999	9905-0140593
PR	23-JUN-1999	9905-0140594
PR	24-JUN-1999	9905-0140695
PR	28-JUN-1999	9905-0140823
PR	29-JUN-1999	9905-0140931
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0142877
PR	01-JUL-1999	9905-0142154
PR	02-JUL-1999	9905-0142055
PR	06-JUL-1999	9905-0142390
PR	08-JUL-1999	9905-0142803
PR	09-JUL-1999	9905-0142820
PR	16-JUL-1999	9905-0144086
PR	19-JUL-1999	9905-0144325
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144334
PR	19-JUL-1999	9905-0144335
PR	20-JUL-1999	9905-0144352
PR	20-JUL-1999	9905-0144632
PR	20-JUL-1999	9905-0144684
PR	21-JUL-1999	9905-0144614
PR	21-JUL-1999	9905-0145016
PR	21-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145192
PR	22-JUL-1999	9905-0145143
PR	23-JUL-1999	9905-0145218
PR	23-JUL-1999	9905-0145524
PR	26-JUL-1999	9905-0145726
PR	27-JUL-1999	9905-0145513
PR	27-JUL-1999	9905-0145519
PR	28-JUL-1999	9905-0145551
PR	02-AUG-1999	9905-0146386
PR	02-AUG-1999	9905-0146388
PR	03-AUG-1999	9905-0146389
PR	04-AUG-1999	9905-0147038
PR	04-AUG-1999	9905-0147302
PR	05-AUG-1999	9905-0147304
PR	05-AUG-1999	9905-0147760
PR	05-AUG-1999	9905-0147920

PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139482.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156559.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159284.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159931.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

|||||:::||||| ||||| || ||
Db 356 MYGKGCFKQPEIKYGDGDDGVNLASLAL-----KVDSLNTVEIDGVSHTSILKDEI 408
Oy 418 ALNEIVGETTSIN 430
|| ||: ||: ||
Db 409 ALKEIMKOISTIN 421

RESULT 8
AAG22980
ID AAG22980 standard; Protein: 413 AA.

XX AAG22980;
AC
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 26115.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
PD
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139891.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142970.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145102.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 58.6%; Score 1367; DB 21; Length 413;
Best Local Similarity 63.1%; Pred. No. 1e-124;
Matches 258; Conservative 58; Mismatches 83; Indels 10; Gaps 4;

Qy 25 LSLTCTCGASNDPLILITGNGNGOLEARLTNOYKRPSTICSEW-YPLIKKKNGWRLWF 83
Db 1 WISMCQAVGSNNYPLILVPGNGNGOLEVRLDREYKRPSSVWCSSWLYPIHKKSGGWRLWF 60

Qy 84 DSSVILAPFTOCFAERMTLHYHOELDDYFNTPGVETRVPHFGSTNSLYLNPRLKHITGY 143
Db 61 DAAVLSPTTRCFSDMMMLTYDPLDDYONAPGVQTRVPHFGSTKSLDLPRLRATSY 120
Qy 144 MAPLVDSLO-KLGADGETLFGAPDYFRYGLAEGHPQSOVSKFLDKNLJEASNSNN 202
Db 121 MEHLVAKALEKKGYVNDQITLGAIFYFRYGLAASGHPRSRAOSFLDDQLQVLEKTSSENE 180
Qy 203 GKPVILLSHSLGGLFVQLNNRNPSPSKKKFIKHFIALSAPMGCAIDEMYTRASGTLY 262
Db 181 GKPVILLSHSLGGLFVQLNNRNPSPSKKKFIKHFIALSAPMGCAIDEMYTRASGTLY 240
Qy 263 PLVDPLLVDEORSSSENNLILPNPKIFGPQ-KPIYTPIRPSAHDMVDPLKDIGEPFG 321
Db 241 PLVNPPLVRHQRTHSENNMLPSTVFHDRTKPLVVTPOVNTATAYEMDRFRADIGFSOG 300
Qy 322 VVPEYTRILPLIGNIKAPOVPIPCINGTGVLLETLYFGKGFDEDEREISYGGDGTVM 381
Db 301 VVPEYKTRVPLPTEBELMTPGVPVTCIYRGVDPEVILMVGKGDFKQPEIKYGDGDTVM 360
Qy 382 VSLIAQSLMKKEKNQYKLVKIDGVSHSIIKDEVALMEIYEITSIN 430
Db 361 ASIALA-----KVDLSMTVEIDGVSHSIIKDEIALKEIMKQIISIN 402

RESULT 9
AAG42647
ID AAG42647 standard; Protein; 413 AA.
XX AAG42647;
AC XX
XX 18-OCT-2000 (first entry)
DT XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 53210.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
PD XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147312.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 58.6%; Score 1367; DB 21; Length 413;
Best Local Similarity 63.1%; Pred. No. 1e-124;
Matches 258; Conservative 58; Mismatches 83; Indels 10; Gaps 4;

QY 25 LSLILCTGASNDPILLINGNGNOLEARLTNOYKSPFICESW-YPLIKKKNGPRLMF 83
DB 1 MMSMGAVGSNNYPLILVNGNGNOLEVRLDREKPSWCSWMLYPIHKSSGWRLMF 60
QY 84 DSSVLIAPPTQCPAERMTLHYHQELDDYFNPQVETRVPHFSGTNSLLYLNPLRIKHITGY 143
DB 61 DAAVILSPPTRCFSDRMMLYDPDLDYQNAFGVQTRVPHFGSTKSLYLDPLRDLATSY 120
QY 144 MAPIVDSIQ-KIGYADGETLFGAPYDFRGYLAEGHPSQVSGSKFLKDLNLFESANSNN 202
DB 121 MEHLVKALEKKKCGYVNDQFIIIGAPYDFRGYLAASGHPNSVASQFLODLQVLEKTSSEME 180
QY 203 GKPVILSHSLGFLVLIQILNENPSPMRKKFIKHFIATLSAPMGCAIDEMYTFASGNTLG 262
DB 181 GKPVILSHSLGFLVLIHNLNRTSPMRKKFIKHFIATLSAPMGCAIDEMYTFASGNTLG 240
QY 263 PLVDPILLVDEQRSSSESNIMLLPNPKIFGPQ-KPIVITPIRYSAHDMVDFLKDIGFPBG 321
DB 241 PLVNPILLVRRHQRTSSNQMLLPSTKVFHDRTKPLVVPQVNTAVEMDRFADIGFSQG 300
QY 322 VVPETRIIPLIGNIKAPVPITCIMGTVGILETILFYKGDPEDEPEISYGDGCTVM 381
DB 301 VVPYKTRVPLTLEELTPGVTCIYRGVDPPEVLMYKGGDKQPEIKYGGDGTVM 360
QY 382 VSLILQSLMKEKKNQYLKVKIKIDGSHRSILKDEVALMEIYGEISIN 430
DB 361 ASLAL-----KVDSLMTVEIDGVSHTSLKDELAKELMKQIITIN 402

RESULT 10
AAG22981
ID AAG22981 standard: Protein; 410 AA.

AC AAG22981;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 26116.

KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PE 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0121820.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.

OS	Arabidopsis thaliana.	
XX		
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000; 2000EP-0301439.	
XX		
PR	25-FEB-1999; 99US-0121825.	PR 30-JUN-1999; 99US-0141287.
PR	05-MAR-1999; 99US-0123180.	PR 01-JUL-1999; 99US-0141842.
PR	09-MAR-1999; 99US-0123548.	PR 01-JUL-1999; 99US-0142154.
PR	23-MAR-1999; 99US-0125788.	PR 02-JUL-1999; 99US-0142055.
PR	25-MAR-1999; 99US-0126264.	PR 06-JUL-1999; 99US-0142390.
PR	29-MAR-1999; 99US-0126785.	PR 08-JUL-1999; 99US-0142803.
PR	01-APR-1999; 99US-0127462.	PR 09-JUL-1999; 99US-0142920.
PR	06-APR-1999; 99US-0128234.	PR 12-JUL-1999; 99US-0142977.
PR	08-APR-1999; 99US-0128714.	PR 13-JUL-1999; 99US-0143542.
PR	16-APR-1999; 99US-0129845.	PR 14-JUL-1999; 99US-0143624.
PR	19-APR-1999; 99US-0130077.	PR 15-JUL-1999; 99US-0144005.
PR	21-APR-1999; 99US-0130449.	PR 16-JUL-1999; 99US-0144085.
PR	23-APR-1999; 99US-0130510.	PR 16-JUL-1999; 99US-0144086.
PR	28-APR-1999; 99US-0130891.	PR 19-JUL-1999; 99US-0144325.
PR	30-APR-1999; 99US-0132048.	PR 19-JUL-1999; 99US-0144331.
PR	30-APR-1999; 99US-0132407.	PR 19-JUL-1999; 99US-0144332.
PR	04-MAY-1999; 99US-0132484.	PR 19-JUL-1999; 99US-0144333.
PR	05-MAY-1999; 99US-0132485.	PR 20-JUL-1999; 99US-0144335.
PR	06-MAY-1999; 99US-0132486.	PR 20-JUL-1999; 99US-0144352.
PR	06-MAY-1999; 99US-0132487.	PR 20-JUL-1999; 99US-0144632.
PR	07-MAY-1999; 99US-0132863.	PR 20-JUL-1999; 99US-0144684.
PR	11-MAY-1999; 99US-0134256.	PR 21-JUL-1999; 99US-0144814.
PR	14-MAY-1999; 99US-0134218.	PR 21-JUL-1999; 99US-0145086.
PR	14-MAY-1999; 99US-0134219.	PR 22-JUL-1999; 99US-0145088.
PR	14-MAY-1999; 99US-0134221.	PR 22-JUL-1999; 99US-0145085.
PR	14-MAY-1999; 99US-0134370.	PR 22-JUL-1999; 99US-0145087.
PR	18-MAY-1999; 99US-0134768.	PR 22-JUL-1999; 99US-0145089.
PR	19-MAY-1999; 99US-0134941.	PR 22-JUL-1999; 99US-0145192.
PR	20-MAY-1999; 99US-0135124.	PR 23-JUL-1999; 99US-0145145.
PR	21-MAY-1999; 99US-0135353.	PR 23-JUL-1999; 99US-0145218.
PR	24-MAY-1999; 99US-0135629.	PR 26-JUL-1999; 99US-0145276.
PR	25-MAY-1999; 99US-0136021.	PR 27-JUL-1999; 99US-0145913.
PR	27-MAY-1999; 99US-0136392.	PR 27-JUL-1999; 99US-0145918.
PR	28-MAY-1999; 99US-0136782.	PR 27-JUL-1999; 99US-0145919.
PR	01-JUN-1999; 99US-0137222.	PR 28-JUL-1999; 99US-0145951.
PR	03-JUN-1999; 99US-0137528.	PR 02-AUG-1999; 99US-0146386.
PR	04-JUN-1999; 99US-0137502.	PR 02-AUG-1999; 99US-0146388.
PR	07-JUN-1999; 99US-0137724.	PR 02-AUG-1999; 99US-0146389.
PR	08-JUN-1999; 99US-0138094.	PR 03-AUG-1999; 99US-0146589.
PR	10-JUN-1999; 99US-0138540.	PR 04-AUG-1999; 99US-0147038.
PR	10-JUN-1999; 99US-0138847.	PR 04-AUG-1999; 99US-0147302.
PR	14-JUN-1999; 99US-0138119.	PR 05-AUG-1999; 99US-0147192.
PR	16-JUN-1999; 99US-0139452.	PR 05-AUG-1999; 99US-0147260.
PR	16-JUN-1999; 99US-0139453.	PR 06-AUG-1999; 99US-0147303.
PR	17-JUN-1999; 99US-0139459.	PR 06-AUG-1999; 99US-0147416.
PR	18-JUN-1999; 99US-0139454.	PR 09-AUG-1999; 99US-0147493.
PR	18-JUN-1999; 99US-0139455.	PR 09-AUG-1999; 99US-0147935.
PR	18-JUN-1999; 99US-0139461.	PR 10-AUG-1999; 99US-0148171.
PR	18-JUN-1999; 99US-0139462.	PR 11-AUG-1999; 99US-0148319.
PR	18-JUN-1999; 99US-0139457.	PR 12-AUG-1999; 99US-0148341.
PR	18-JUN-1999; 99US-0139458.	PR 13-AUG-1999; 99US-0148565.
PR	18-JUN-1999; 99US-0139459.	PR 13-AUG-1999; 99US-0148684.
PR	18-JUN-1999; 99US-0139460.	PR 16-AUG-1999; 99US-0149368.
PR	18-JUN-1999; 99US-0139461.	PR 17-AUG-1999; 99US-0149375.
PR	18-JUN-1999; 99US-0139462.	PR 18-AUG-1999; 99US-0149426.
PR	18-JUN-1999; 99US-0139463.	PR 20-AUG-1999; 99US-0149722.
PR	18-JUN-1999; 99US-0139750.	PR 20-AUG-1999; 99US-0149723.
PR	18-JUN-1999; 99US-0139763.	PR 20-AUG-1999; 99US-0149929.
PR	21-JUN-1999; 99US-0139817.	PR 23-AUG-1999; 99US-0149902.
PR	22-JUN-1999; 99US-0139899.	PR 23-AUG-1999; 99US-0149930.
PR	23-JUN-1999; 99US-0140353.	PR 25-AUG-1999; 99US-0150566.
PR	23-JUN-1999; 99US-0140354.	PR 26-AUG-1999; 99US-0150584.
PR	24-JUN-1999; 99US-0140695.	PR 27-AUG-1999; 99US-0151065.
PR	28-JUN-1999; 99US-0140823.	PR 27-AUG-1999; 99US-0151066.
PR	29-JUN-1999; 99US-0140991.	PR 30-AUG-1999; 99US-0151080.
		PR 30-AUG-1999; 99US-0151303.
		PR 31-AUG-1999; 99US-0151338.
		PR 01-SEP-1999; 99US-0151930.
		PR 07-SEP-1999; 99US-0152563.
		PR 10-SEP-1999; 99US-0153070.
		PR 13-SEP-1999; 99US-0153758.

```
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 58.6%; Score 1366; DB 21; Length 410;
Best Local Similarity 63.5%; Pred. No. 1.3e-124;
Matches 258; Conservative 56; Mismatches 82; Indels 10; Gaps 4;
```

```
QY 28 LCTGASNLDPILIPGNGNQLRLNQQYKPSFTICEBW-YPLIKKNGFRLMFDSS 86
DB 1 MCOAVGSINYYPLILVPGNGNQLRLDREYKPPSSVWCSSMLYPIHKKSGWFLMFDNA 60
QY 87 VTLAFTGCFARMTLHYHQELDDYFNTPGVETRVPHGSTNSLLYLNRLKHITGYMAP 146
DB 61 VLLSPFTKCFSDRMMLYDDPDLDDYQNAFGVQTVRPHFGSTKSLLYLDRLDATSYMEH 120
QY 147 LVDSHQ-KLGYADGETLFGAPYDFRYGLAEGHPQVGSCKFLDKLNLIEFASNSNGKP 205
DB 121 LVKALEKKCGYVNDQTIIIGAPYDFRYGLAEGHPQVGSCKFLDKLNLIEFASNSNECKP 180
QY 206 VILLSHSLGGFLVQLLNRPNSMRKKFKIHFIALSAPMGAIIDEMTYFASGNTLGVPV 265
DB 181 VILLSHSLGGFLVQLLNRPNSMRKKFKIHFIALSAPMGAIIDEMTYFASGNTLGVPV 240
QY 266 DPLVLRDQSRSSNLMLLPNPKITGPQ-KPIVITPIRYSAMDVDELKIDGFEQYIP 324
DB 241 NPLLRHRRQSRSSNLMLLPNPKITGPQ-KPIVITPIRYSAMDVDELKIDGFEQYIP 300
QY 325 YETRIPLIGNIKARPOVITICMGVGTLETIFYGKGDDEPRISTSGDGDGTIVNLSL 384
DB 301 YKTRPLPLTEELMPTGCVPTCTCYGRGVDTPEVYLMGKGSGFDPQPIKIDGSGGTIVNLSL 360
QY 385 LALQSLMEKKENQYLVKVDGVSHTSILKDEVALNEIYGETITSIN 430
```

```
DB 361 AAL-----KYDSINTWEIDGVSHTSILKDEIALKEIMKQIISIN 399
RESULT 12
AAB24263
ID AAB24263 standard; Protein; 387 AA.
XX
XX AAB24263;
AC
XX 12-FEB-2001 (first entry)
XX
XX Arabidopsis thaliana PDAT amino acid sequence SEQ ID NO:2a.
DE
XX PDAT; phospholipid:diacylglycerol acyltransferase; triacylglycerol;
KW TAG; EST; expressed sequence tag; fatty acid; oil content.
XX
XX Arabidopsis thaliana.
OS
XX WO200060095-A2.
PN
XX 12-OCT-2000.
PD
XX
XX 28-MAR-2000; 2000WO-EP02701.
PF
XX 01-APR-1999; 99EP-0106656.
PR 10-JUN-1999; 99EP-0111321.
PR 07-FEB-2000; 2000US-0180687.
XX
XX (BADI ) BASF PLANT SCI GMBH.
XX
XX Dahlqvist A, Stahl U, Lenman M, Banas A, Ronne H, Slymne S;
PI WPI; 2000-665012/64.
XX
XX Phospholipid:diacylglycerol acyltransferase enzymes in the biosynthetic
PT pathway for triacylglycerol production and DMS encoding them, useful
PT for producing triacylglycerol, or for transforming any cell or organism
XX to increase oil content -
XX
XX Claim 5; Page 68-69; 97pp; English.
XX
XX The present invention describes an enzyme for catalysing (in an
CC acyl-CoA-independent reaction) the transfer of fatty acids from
CC phospholipids to diacylglycerol in the biosynthetic pathway for the
CC production of triacylglycerol (TAG). The enzyme is designated as
CC phospholipid:diacylglycerol acyltransferase (PDAT). The enzyme and
CC the nucleotides encoding them are useful for producing TAG and/or
CC TAG with uncommon fatty acids. The enzyme and the nucleotide are
CC also useful for transforming any cell or organism in order to be
CC expressed in this cell or organism and result in an altered, preferably
CC increased oil content of this cell or organism. The present sequence
CC represents an Arabidopsis thaliana PDAT amino acid sequence.
XX
XX Sequence 387 AA;
SQ
Query Match 57.7%; Score 1345; DB 21; Length 387;
Best Local Similarity 64.8%; Pred. No. 1.3e-122;
Matches 254; Conservative 52; Mismatches 76; Indels 10; Gaps 4;
```

79 FRLWEDSSVIL-APFTQCFAERMTHYHQELDDYFNTPGVETRVPHFGSTNSLLYNPRL 137

Query Match	Score	DB	Length
31.78;	739.5;	21;	439;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 19:13:35 ; Search time 17 Seconds
(without alignments)
766.727 Million cell updates/sec

Title: US-09-857-612A-14

Perfect score: 2333
Sequence: 1 MKKEQEGGLAEVATLTFTV.....GETTSINSHAEGLSINLPSG 443

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCtus.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106.5	4.6	314	US-09-355-166-17	Sequence 17, Appl
2	98.5	4.2	566	US-08-104-158-2	Sequence 2, Appl
3	97.5	4.2	346	US-08-602-359A-34	Sequence 34, Appl
4	97.5	4.2	444	US-08-485-938A-33	Sequence 33, Appl
5	95.5	4.1	3169	US-09-453-702B-257	Sequence 257, App
6	92.5	4.0	602	US-08-446-100-10	Sequence 10, Appl
7	92	3.9	444	US-08-483-140-28	Sequence 28, Appl
8	92	3.9	444	US-08-485-938A-32	Sequence 32, Appl
9	90.5	3.9	685	US-08-947-965-74	Sequence 74, Appl
10	90.5	3.9	1429	US-07-642-002-2	Sequence 2, Appl
11	90.5	3.9	1429	US-08-365-486A-13	Sequence 13, Appl
12	90.5	3.9	1429	US-08-319-866-11	Sequence 11, Appl
13	90.5	3.9	1429	US-08-880-342-13	Sequence 13, Appl
14	90.5	3.9	1430	US-08-705-625-4	Sequence 4, Appl
15	90.5	3.9	1430	US-09-010-998-5	Sequence 5, Appl
16	90.5	3.9	1430	US-09-022-574-4	Sequence 4, Appl
17	89.5	3.8	435	US-09-072-917A-9	Sequence 9, Appl
18	89.5	3.8	602	US-08-446-100-9	Sequence 9, Appl
19	89.5	3.8	602	US-08-446-100-11	Sequence 11, Appl
20	89.5	3.8	2756	US-08-375-709-11	Sequence 11, Appl
21	89.5	3.8	2756	US-08-752-929-11	Sequence 11, Appl
22	89.5	3.8	2756	US-09-090-793-7	Sequence 7, Appl
23	89	3.8	501	US-08-906-791-2	Sequence 2, Appl
24	88.5	3.8	602	US-08-446-100-5	Sequence 5, Appl
25	88.5	3.8	602	US-08-446-100-12	Sequence 12, Appl
26	87.5	3.8	526	US-08-852-401-3	Sequence 3, Appl
27	87.5	3.8	602	US-08-446-100-6	Sequence 6, Appl

28	87.5	3.8	602	US-08-446-100-15	Sequence 15, Appl
29	87.5	3.8	602	US-08-446-100-16	Sequence 16, Appl
30	87.5	3.8	660	US-09-134-001C-3350	Sequence 3350, Ap
31	87	3.7	251	US-09-345-469-4	Sequence 4, Appl
32	87	3.7	501	US-09-111-730-1	Sequence 1, Appl
33	87	3.7	624	US-09-877-730-24	Sequence 24, Appl
34	86.5	3.7	602	US-08-446-100-1	Sequence 1, Appl
35	86.5	3.7	602	US-08-446-100-2	Sequence 2, Appl
36	86.5	3.7	602	US-08-446-100-3	Sequence 3, Appl
37	86.5	3.7	602	US-08-446-100-4	Sequence 4, Appl
38	86.5	3.7	602	US-08-446-100-8	Sequence 8, Appl
39	86.5	3.7	602	US-08-446-100-18	Sequence 18, Appl
40	86.5	3.7	602	US-08-446-100-24	Sequence 24, Appl
41	86.5	3.7	602	US-09-334-489-3	Sequence 3, Appl
42	86.5	3.7	602	US-09-334-489-4	Sequence 4, Appl
43	86	3.7	933	US-08-313-200-1	Sequence 1, Appl
44	86	3.7	933	PCW-US93-03837-1	Sequence 1, Appl
45	85.5	3.7	394	US-07-603-133B-11	Sequence 11, Appl

ALIGNMENTS

```
RESULT 1
US-09-355-166-17
; Sequence 17, Application US/09355166
; Patent No. 6316241
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
; FILE REFERENCE: GC511-PCR
; CURRENT APPLICATION NUMBER: US/09/355,166
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Bacillus
; US-09-355-166-17

Query Match          4.6%; Score 106.5; DB 4; Length 314;
Best Local Similarity 21.9%; Pred. No. 0.0096;
Matches 65; Conservative 34; Mismatches 109; Indels 89; Gaps 12;

QY 71 LKRRNGWRLWFDSSVILAPFCFAERMTLHYHOELDDYFNTPGVETRPVH----- 123
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 4 LVDSRNHYTLRFS--LRGMSYCMKEQOTDRTNGGTSMNFTIPGTEVRRMSRNERT 60
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 124 -----FGSTNSLLYLNPRKHTTGYPAPLVDSLOKIG 155
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 61 YHIFISKPSTPPPPAGPYVYILLDANSVGTWTEAVRIQGRPEKTVPAIVG--IG 117
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 156 YADGETLFCAPY-DPRYGLAEGHPSOVGS-----KFL-KDKMLIEASNS 200
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 118 VETAEPRASRRKRDTPMPAQSCLPERPPGRMPHGAEGFRFIEELKKEIERDYI 177
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 201 NNGKPYVILLSHLSGLLEVLQLLNRP-----PS--WRKKFL---KHFIALSAW 244
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 178 DKRQGTI-FGHSLGLEVLQVLLTKPDARQTYIAGSPSTHMKPFLIKTKTDHFVSLTKN 236
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 245 GCAID-----EMTFASGNTLGVPLVPLLVNDEQR-----SESIMLLP 285
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 237 NOPINLLAAGLELHHSKSRMNDNARELYERLAVLSEQIRAFCEFGEGHISVLP 293
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 2
; Sequence 2, Application US/08104158
; Patent No. 6215042
; GENERAL INFORMATION:
; APPLICANT: Willmitzer, Lothar
; APPLICANT: Sonnewald, Uwe
```


Db 255 GYGRMLAGVLDGFGPAGYGRVDKPLLVYGRPRDPLVTDEARSLAS 301

RESULT 4

US-08-485-938A-33

Sequence 33, Application US/08485938A

Patent No. 5847088

GENERAL INFORMATION:

APPLICANT: Cousens, Lawrence S.

APPLICANT: Eberhardt, Christine D.

APPLICANT: Gray, Patrick W.

APPLICANT: Le Trong, Hai

APPLICANT: Toelker, Larry W.

APPLICANT: Wilder, Cheryl L.

TITLE OF INVENTION: Platelet-Activating Factor

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,938A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/318,905

FILING DATE: 06-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/133,803

FILING DATE: 06-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: No. 5847088and, Greta E.

REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 27866/32792

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 474-6300

TELEFAX: (312) 474-0448

TELEX: 25-3658

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 444 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-485-938A-33

Query Match 4.2%; Score 97.5; DB 2; Length 444;

Best Local Similarity 22.4%; Pred. No. 0.15;

Matches 54; Conservative 39; Mismatches 85; Indels 63; Gaps 10;

Db 378 SNVAISLSNKASIAFLQKHILGLOKDFDQWDSLVESGDHNLIPOTNINTNHQAALLQNSRG 437

Qy 419 L 419

Db 438 I 438

RESULT 5

US-09-453-702B-257

Sequence 257, Application US/09453702B

Patent No. 6365723

GENERAL INFORMATION:

APPLICANT: Blattner, Frederick R.

APPLICANT: Burland, Valerie

APPLICANT: Perna, Nicole T.

APPLICANT: Plunkett, Guy

APPLICANT: Welch, Rod

TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESS: Quarles & Brady

STREET: 1 South Pinckney Street

CITY: Madison

STATE: WI

COUNTRY: US

ZIP: 53701-2113

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 8.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/453,702B

FILING DATE: 03-Dec-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/110,955

FILING DATE: 04-Dec-1998

ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J.

REGISTRATION NUMBER: 27386

REFERENCE/DOCKET NUMBER: 960296, 95017

TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 251-5000

TELEFAX: (608) 251-9166

INFORMATION FOR SEQ ID NO: 257:

SEQUENCE CHARACTERISTICS:

LENGTH: 3169 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 257;

US-09-453-702B-257

Query Match 4.1%; Score 95.5; DB 4; Length 3169;

Best Local Similarity 19.3%; Pred. No. 6.2;

Matches 87; Conservative 62; Mismatches 152; Indels 149; Gaps 23;

Db 189 DLKNIIEASNSNNGKPYVLLSHSLGGLFVLQLLNRPSPMKKFIKHLIALSA---PMG 245

Qy 245 DLKNIIEASNSNNGKPYVLLSHSLGGLFVLQLLNRPSPMKKFIKHLIALSA---PMG 245

Db 251 DLFPDEQAKSDIRDKIAIIGHISFGATVIOIQLSED---QRF-RGCIADLDMMPVG 304

Qy 246 GAIDEMTFASGNTGLVPLVDLVLVDEQRSSSNMLLPNPKIFGPKPIVITPIRYS 305

Db 305 ---DEVY-----SRIPQLP---FFINSERFQYPSN---IIRMKCFIPDR----- 340

Qy 306 AHDWVFLKIDIFPGGVVPEYTRILPLIGNIKAPQVPTICMGTVGLTLETFYGGDPD 365

Db 341 -----ERKMITIRGSHQNFVDYFPAFSKIGLYFTL---KGIDID 377

Qy 366 ERPEISYGGDGTVMVLSIALQ-----SLMKEEKNOYLAVKIKIDGVSHSILKDEVA 418

Db 189 DLKNIIEASNSNNGKPYVLLSHSLGGLFVLQLLNRPSPMKKFIKHLIALSA---PMG 245

Qy 245 DLKNIIEASNSNNGKPYVLLSHSLGGLFVLQLLNRPSPMKKFIKHLIALSA---PMG 245

Db 251 DLFPDEQAKSDIRDKIAIIGHISFGATVIOIQLSED---QRF-RGCIADLDMMPVG 304

Qy 246 GAIDEMTFASGNTGLVPLVDLVLVDEQRSSSNMLLPNPKIFGPKPIVITPIRYS 305

Db 305 ---DEVY-----SRIPQLP---FFINSERFQYPSN---IIRMKCFIPDR----- 340

Qy 306 AHDWVFLKIDIFPGGVVPEYTRILPLIGNIKAPQVPTICMGTVGLTLETFYGGDPD 365

Db 341 -----ERKMITIRGSHQNFVDYFPAFSKIGLYFTL---KGIDID 377

Qy 366 ERPEISYGGDGTVMVLSIALQ-----SLMKEEKNOYLAVKIKIDGVSHSILKDEVA 418

Db 189 DLKNIIEASNSNNGKPYVLLSHSLGGLFVLQLLNRPSPMKKFIKHLIALSA---PMG 245

Qy 245 DLKNIIEASNSNNGKPYVLLSHSLGGLFVLQLLNRPSPMKKFIKHLIALSA---PMG 245

Db 251 DLFPDEQAKSDIRDKIAIIGHISFGATVIOIQLSED---QRF-RGCIADLDMMPVG 304

Qy 246 GAIDEMTFASGNTGLVPLVDLVLVDEQRSSSNMLLPNPKIFGPKPIVITPIRYS 305

Db 305 ---DEVY-----SRIPQLP---FFINSERFQYPSN---IIRMKCFIPDR----- 340

Qy 306 AHDWVFLKIDIFPGGVVPEYTRILPLIGNIKAPQVPTICMGTVGLTLETFYGGDPD 365

Db 341 -----ERKMITIRGSHQNFVDYFPAFSKIGLYFTL---KGIDID 377

Qy 366 ERPEISYGGDGTVMVLSIALQ-----SLMKEEKNOYLAVKIKIDGVSHSILKDEVA 418

Db 189 DLKNIIEASNSNNGKPYVLLSHSLGGLFVLQLLNRPSPMKKFIKHLIALSA---PMG 245

Qy 245 DLKNIIEASNSNNGKPYVLLSHSLGGLFVLQLLNRPSPMKKFIKHLIALSA---PMG 245

Db 251 DLFPDEQAKSDIRDKIAIIGHISFGATVIOIQLSED---QRF-RGCIADLDMMPVG 304

Qy 246 GAIDEMTFASGNTGLVPLVDLVLVDEQRSSSNMLLPNPKIFGPKPIVITPIRYS 305

Db 305 ---DEVY-----SRIPQLP---FFINSERFQYPSN---IIRMKCFIPDR----- 340

Qy 306 AHDWVFLKIDIFPGGVVPEYTRILPLIGNIKAPQVPTICMGTVGLTLETFYGGDPD 365

Db 341 -----ERKMITIRGSHQNFVDYFPAFSKIGLYFTL---KGIDID 377

Qy 366 ERPEISYGGDGTVMVLSIALQ-----SLMKEEKNOYLAVKIKIDGVSHSILKDEVA 418

Db 189 DLKNIIEASNSNNGKPYVLLSHSLGGLFVLQLLNRPSPMKKFIKHLIALSA---PMG 245

Query Match 3.9%: Score 92; DB 1; Length 444;
 Best Local Similarity 20.5%: Pred. No. 0.58;
 Matches 98; Conservative 51; Mismatches 122; Indels 208; Gaps 26;

```

QY 26 SLCTCGA-----SNLDPLILPGNG-GNQLERLTNOYKPTFTCESWYPLIKK 75
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 8 ALFCLSCILVHPIDMODLNPVANHRSAMAKIALMA-----AASIRQSRIP---KG 59
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 76 NGFR-----LWFDSSVILAPFTQCAERMT---LHYHQLD-----110
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 60 NGSYVGCTDLMFD-----YTNKGTFLRLYPSQEDHSDTLWIPNKEFFGLS 108
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 111 -YENFTGVERVP--HFGSNSLLYLNPRLKHTGYMAPLYDSLOKLG-----155
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 109 KILGTWLMKILSFFFGSYTTPANNNSPLR--TGEKYPLIVSHGIGAFRTIYSAIGID 166
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 156 -----YADGETLFGAPYDFRYGLAAGHPSOVGSK---FLKDLK-----191
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 167 LASHGFIYVAIEHRDGA--SATYFFKDSAAE-----IGNKMSVYLQELKPGDEDEIHYR 219
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 192 -----NLI-----EASNSNNGKPVILLSHSLGLF 217
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 220 NEQVORAKESQALNLIIDIDHGRPIKNVLDLEFVDEQLKDISIDRKIAVIGHSFGAT 279
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 218 VIQLLRNPPSMRKKFKIHFIASAPWGGAIDEMTFASGNTLGVPPLVPLVRDQRSS 277
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 280 VLOALSD-----QRF-KCIGALDA-W-----MLPLDAIYSR-----310
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 278 ESNMLLPNPKIFGPQKPIYTPIRPYSAHDWVFLKDIGFPEGVVPY-----ETRI 329
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 311 -----IPQPLF-----INSER-----FQFPENIKMKKCYSPDKERKM 344
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 330 LPLIGNIKAPQVYITCIMGVGTLETFYKGKDFDERPEISYGDGDGYVNLVSLALQ 388
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 345 ITRGSVHONFADFTTGTGIVGIYFTL---KGDIDSNVAID-----LCNKASIAFLQ 394
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 8
 US-08-485-938A-32
 ; Sequence 32, Application US/08485938A
 ; Patent No. 5847088
 ; GENERAL INFORMATION:
 ; APPLICANT: Cousens, Lawrence S.
 ; APPLICANT: Eberhardt, Christine D.
 ; APPLICANT: Gray, Patrick W.
 ; APPLICANT: Le Trong, Hai
 ; APPLICANT: Tjoelker, Larry W.
 ; APPLICANT: Wilder, Cheryl L.
 ; TITLE OF INVENTION: Platelet-activating Factor
 ; NUMBER OF SEQUENCES: 36
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION NUMBER: US/08/485,938A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/318,905
 ; FILING DATE: 06-OCT-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/133,803

FILING DATE: 06-OCT-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5847088and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 27866/32792
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEX: 25-3658
 ; INFORMATION FOR SEQ ID NO: 32:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 444 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-485-938A-32

Query Match 3.9%: Score 92; DB 2; Length 444;
 Best Local Similarity 20.5%: Pred. No. 0.58;
 Matches 98; Conservative 51; Mismatches 122; Indels 208; Gaps 26;

```

QY 26 SLCTCGA-----SNLDPLILPGNG-GNQLERLTNOYKPTFTCESWYPLIKK 75
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 8 ALFCLSCILVHPIDMODLNPVANHRSAMAKIALMA-----AASIRQSRIP---KG 59
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 76 NGFR-----LWFDSSVILAPFTQCAERMT---LHYHQLD-----110
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 60 NGSYVGCTDLMFD-----YTNKGTFLRLYPSQEDHSDTLWIPNKEFFGLS 108
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 111 -YENFTGVERVP--HFGSNSLLYLNPRLKHTGYMAPLYDSLOKLG-----155
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 109 KILGTWLMKILSFFFGSYTTPANNNSPLR--TGEKYPLIVSHGIGAFRTIYSAIGID 166
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 156 -----YADGETLFGAPYDFRYGLAAGHPSOVGSK---FLKDLK-----191
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 167 LASHGFIYVAIEHRDGA--SATYFFKDSAAE-----IGNKMSVYLQELKPGDEDEIHYR 219
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 192 -----NLI-----EASNSNNGKPVILLSHSLGLF 217
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 220 NEQVORAKESQALNLIIDIDHGRPIKNVLDLEFVDEQLKDISIDRKIAVIGHSFGAT 279
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 218 VIQLLRNPPSMRKKFKIHFIASAPWGGAIDEMTFASGNTLGVPPLVPLVRDQRSS 277
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 280 VLOALSD-----QRF-KCIGALDA-W-----MLPLDAIYSR-----310
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 278 ESNMLLPNPKIFGPQKPIYTPIRPYSAHDWVFLKDIGFPEGVVPY-----ETRI 329
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 311 -----IPQPLF-----INSER-----FQFPENIKMKKCYSPDKERKM 344
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 330 LPLIGNIKAPQVYITCIMGVGTLETFYKGKDFDERPEISYGDGDGYVNLVSLALQ 388
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 345 ITRGSVHONFADFTTGTGIVGIYFTL---KGDIDSNVAID-----LCNKASIAFLQ 394
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 9
 US-08-947-965-74
 ; Sequence 74, Application US/08947965A
 ; Patent No. 6004790
 ; GENERAL INFORMATION:
 ; APPLICANT: Dijkhuizen, Lubbert
 ; APPLICANT: Andersen, Carsten
 ; APPLICANT: Osten, Claus von der
 ; TITLE OF INVENTION: Cyclomaltodextrin Glucanotransferase
 ; FILE REFERENCE: 4285.204-US
 ; CURRENT APPLICATION NUMBER: US/08/947,965A
 ; FILING DATE: 1997-10-09
 ; EARLIER APPLICATION NUMBER: 0477/95
 ; EARLIER FILING DATE: 1995-04-21
 ; EARLIER APPLICATION NUMBER: 1173/95
 ; EARLIER FILING DATE: 1995-10-17
 ; EARLIER APPLICATION NUMBER: 1281/95


```
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,342
FILING DATE: 23-JUN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB95/00996
FILING DATE: 13-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/365,486
FILING DATE: 23-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8255-0018.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1429 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-880-342-13
```

```
Query Match          3.9%; Score 90.5; DB 4; Length 1429;
Best Local Similarity 18.5%; Pred. No. 5.7;
Matches 67; Conservative 58; Mismatches 142; Indels 95; Gaps 16;
```

```
QY 108 LDDYFNTPGVETRPVPHGSGTNSLLYLNPRLKHTGYAPRLVDSLOKLGADGETLFGAPY 167
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 362 LDQYSS-----IKRFGSKAHM---DRLEEVN---KELESTYQLKDTLETIGAKH 407
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 168 DFRYGLAEG-----HPSQVSKFLDKLKNLIEASNSNGKRPVI-LLSHSLG 214
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 408 AMRNASRCVGRIOWSKIQVDPARDCTTAHGMFNVCNHVATNKGKLBRSALITFPQRTD 467
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 215 GLEVLQLLNRPSPWRKKFKHFTALSPWGAIDEMTTASGNTLGAPL---VDPILVR 271
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 468 GKHDFFV-----WNSQLIRY-----AGYKQPDSTLGDPRANVOFTELCIQ 507
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 272 DEORSSSENMLLP-----NPKIRG-PQKPIVITPIRPSAHMDVDFLKDIGPPEGV 322
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 508 QGKAPRGRDVLPLLIQANGNDPELFQIRPELVLEVPPIR---HPRFDMFKDLGLKMWG 563
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 323 YPEETRIPLIGNIKARQVPIIT-CIMGTGVTLETLEFGKDPDEREPEISGDGDGTVNL 381
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 564 LPVAVSNMLLEIGLEFACPSFGWMTETIGVRD-----YCD-NSRYNI 606
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 382 VSLA-----LQSLMKEEKQYLYKIDGVSHTSILKDEVALNEIVGETTSINSH 432
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 607 LEEVAKKMDLDMRTSSLMKDO-----ALVEINIAVLVSQSDKVTIVDHSATESPTKH 661
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 433 AE 434
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 662 ME 663
```

```
RESULT 14
; Sequence 4, Application US/08705625
; Patent No. 5908756
; GENERAL INFORMATION:
; APPLICANT: Snyder, Solomon H.
; APPLICANT: Jaffrey, Samle R.
; TITLE OF INVENTION: Protein Inhibitor of Neuronal Nitric
; Oxide Synthase
```

```
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
COUNTRY: U.S.A.
ZIP: 20001-4597
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/705,625
FILING DATE: 30-AUG-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.57071
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9299
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1430 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-705-625-4
```

```
Query Match          3.9%; Score 90.5; DB 2; Length 1430;
Best Local Similarity 18.5%; Pred. No. 5.7;
Matches 67; Conservative 58; Mismatches 142; Indels 95; Gaps 16;
```

```
QY 108 LDDYFNTPGVETRPVPHGSGTNSLLYLNPRLKHTGYAPRLVDSLOKLGADGETLFGAPY 167
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 362 LDQYSS-----IKRFGSKAHM---DRLEEVN---KELESTYQLKDTLETIGAKH 407
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 168 DFRYGLAEG-----HPSQVSKFLDKLKNLIEASNSNGKRPVI-LLSHSLG 214
    :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
DB 408 AMRNASRCVGRIOWSKIQVDPARDCTTAHGMFNVCNHVATNKGKLBRSALITFPQRTD 467
    :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~
QY 215 GLEVLQLLNRPSPWRKKFKHFTALSPWGAIDEMTTASGNTLGAPL---VDPILVR 271
    :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~
DB 468 GKHDFFV-----WNSQLIRY-----AGYKQPDSTLGDPRANVOFTELCIQ 507
    :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~
QY 272 DEORSSSENMLLP-----NPKIRG-PQKPIVITPIRPSAHMDVDFLKDIGPPEGV 322
    :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~
DB 508 QGKAPRGRDVLPLLIQANGNDPELFQIRPELVLEVPPIR---HPRFDMFKDLGLKMWG 563
    :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~
QY 323 YPEETRIPLIGNIKARQVPIIT-CIMGTGVTLETLEFGKDPDEREPEISGDGDGTVNL 381
    :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~
DB 564 LPVAVSNMLLEIGLEFACPSFGWMTETIGVRD-----YCD-NSRYNI 606
    :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~
QY 382 VSLA-----LQSLMKEEKQYLYKIDGVSHTSILKDEVALNEIVGETTSINSH 432
    :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~
DB 607 LEEVAKKMDLDMRTSSLMKDO-----ALVEINIAVLVSQSDKVTIVDHSATESPTKH 661
    :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~
QY 433 AE 434
    :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~
DB 662 ME 663
```

```
RESULT 15
; Sequence 5, Application US/09010998
; Patent No. 6103872
; GENERAL INFORMATION:
; APPLICANT: Snyder, Solomon
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 19:15:26 ; Search time 30 Seconds
(without alignments)
377.272 Million cell updates/sec

Title: US-09-857-612A-14
Perfect score: 2333
Sequence: 1 MKKEQEGELKIEVATLTVTY.....GEITSINSHAEIGLSNLFSG 443

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEM_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCr_NEM_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEM_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEM_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCrUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEM_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEM_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEM_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	610	26.1	412	9	US-09-978-295A-157
2	610	26.1	412	9	US-09-978-697-157
3	610	26.1	412	9	US-09-978-192A-157
4	610	26.1	412	9	US-09-999-812A-157
5	610	26.1	412	9	US-10-001-054-8
6	610	26.1	412	9	US-09-978-189-157
7	541	23.2	440	10	US-09-919-497-80
8	112	4.8	300	9	US-09-738-626-3594
9	106.5	4.6	314	10	US-09-950-368-17
10	103.5	4.4	338	9	US-09-738-626-5870
11	101.5	4.4	574	10	US-09-748-739A-22
12	100.5	4.3	342	10	US-09-896-578-2
13	97.5	4.2	346	10	US-10-027-805-34
14	97.5	4.2	346	10	US-09-903-410-34
15	97.5	4.2	1014	9	US-09-738-626-4656
16	97	4.2	338	9	US-09-738-626-5692
17	95.5	4.1	3169	9	US-10-114-170-257
18	91	3.9	389	10	US-09-815-242-10411
19	90.5	3.9	574	10	US-09-748-739A-6

20	89.5	3.8	416	10	US-09-731-231A-6	Sequence 6, Appl1
21	89.5	3.8	618	9	US-09-738-626-5766	Sequence 5766, Ap
22	89.5	3.8	7968	9	US-10-077-110-5	Sequence 5, Appl1
23	89	3.8	350	10	US-09-815-242-5749	Sequence 5749, Ap
24	89	3.8	350	10	US-09-815-242-12468	Sequence 12468, A
25	89	3.8	829	10	US-09-801-368-324	Sequence 324, Appl
26	87.5	3.8	333	10	US-09-794-715A-16	Sequence 16, Appl
27	87	3.7	233	9	US-10-056-744B-4	Sequence 4, Appl1
28	87	3.7	251	9	US-10-068-134-4	Sequence 4, Appl1
29	87	3.7	501	10	US-09-878-262B-1	Sequence 1, Appl1
30	87	3.7	624	9	US-10-231-353-24	Sequence 24, Appl
31	86.5	3.7	574	10	US-09-748-739A-4	Sequence 4, Appl1
32	86.5	3.7	574	10	US-09-748-739A-8	Sequence 8, Appl1
33	86.5	3.7	574	10	US-09-748-739A-17	Sequence 17, Appl
34	86.5	3.7	574	10	US-09-748-739A-19	Sequence 19, Appl
35	86.5	3.7	574	10	US-09-748-739A-20	Sequence 20, Appl
36	86.5	3.7	602	10	US-09-748-739A-2	Sequence 2, Appl1
37	84.5	3.6	811	9	US-09-738-626-5817	Sequence 5817, Ap
38	84	3.6	574	10	US-09-748-739A-18	Sequence 18, Appl
39	83.5	3.6	811	10	US-09-895-382-32	Sequence 32, Appl
40	83.5	3.6	1399	10	US-09-815-242-5179	Sequence 5179, Ap
41	83.5	3.6	1610	9	US-09-738-626-6666	Sequence 6666, Ap
42	83	3.6	322	9	US-09-738-626-3508	Sequence 3508, Ap
43	83	3.6	477	9	US-10-047-412A-10	Sequence 10, Appl
44	83	3.6	477	10	US-09-923-556-6	Sequence 6, Appl1
45	83	3.6	477	10	US-09-987-025-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-978-295A-157
; Sequence 157, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C11
; CURRENT APPLICATION NUMBER: US/09/978, 295A
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249

[illegible]


```

; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match          26.1%; Score 610; DB 9; Length 412;
Best Local Similarity 34.8%; Pred. No. 2e-49;
Matches 141; Conservative 80; Mismatches 156; Indels 28; Gaps 12;

QY 21 VVYMLSLC--TCGASNDPLILIPNGSGNQLBARLTNOVKPST--FICGSWPLIKKN 76
DB 20 LLLLLMLADPALPAGRHPPVVLPGDLGNQLAKLD--KPYVHYLCS-----KTE 70
QY 77 GWFRLMFSSVILAPFTQCAFEMTLHYHOELDVFNTPGVETRVPHFGSTNSLYLNP 136
DB 71 SYFTIMLNLELLPVLIIDCMIDNIRLYNKTSRATQFPDGVDRVPGFGKTFLEFLDPS 130
QY 137 LKHTTYMAPLVDSLQKLYADGETLFGADYDFRYGIAAGHPSQVGSFKLDKLNLEE 196
DB 131 KSSGVSEYFHMYVESLWGMVGTGCEDEVGAPYDWRA-----PNEGPFYLA-LREIMEE 183
QY 197 ASSSNGKPVILSHSLGGLFVQLNLRNPPSRKKFIKFIHLSAPMGAIDEMTEFAS 256
DB 184 MYQLYGG-PVVLVAHSMGNMYTYLFLQROQAWKDYIRAFVSLGAPWGVAATLRLVAS 242
QY 257 GNTLGPVLVDLIVRDEORSSESMLMLPNPKJFGPKPIVPIPIRPSAHDMVFLKDI 316
DB 243 GDNNRIPVIGLTKIRQGRSAVSTSWMLPYNTWSPKRYVQPTTYTLADRKRFQDI 302
QY 317 GPPEG-VPEYETRIPLIGNIKAPQVPYTCIMGTGVTLETLFYGKGDFDER-PEISYGD 374
DB 303 GFEDGMLMRDTE--GLVEATMPPGVQLHCLYGTGVPTPSPFY--ESFPDRPKICFGD 358
QY 375 GDSGVMLVSLALOSLMKEKNQYLVKVKIDGVSHTSILKDEVAL 419
DB 359 GDSGVMLKSLAQCA-WQSRQHOVLLOELPGSEHIEMLNATTL 402

RESULT 2
US-09-978-697-157
; Sequence 157, Application US/09978697
; Patent No. US20020169284A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurley, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kijavitt, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
```

```

; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630PIC27
; CURRENT APPLICATION NUMBER: US/09/978,697
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
```

;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07

;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 26.1%; Score 610; DB 9; Length 412;
Best Local Similarity 34.8%; Pred. No. 2e-49;
Matches 141; Conservative 80; Mismatches 156; Indels 28; Gaps 12;

Qy 21 VVWMLSLC--TCGASNDPILIPGNGNQLERLTQNYKPRST--FICESWYPLIKKN 76
Db 20 LLLILMLADPALPAGRPVPLVPGDLGNQLEAKLD--KPTVVHYLCS-----KTE 70
Qy 77 GWERLWFDSSVILAFPTCCFAERMTHHQELDDYFNTPGVETKVRPHNGSINSLYLNPR 136
Db 71 SYFTIMLNELELLPYIIDCWIDNIRLVYNNKTSRATQPPDGDVVRVPGKTFSSLEFLDPS 130
Qy 137 LKHITGYAAPLVDSLOKGYADGETLFGAPDFERYGLAEGHPSOVGSKFLDKNLTEE 196
Db 131 KSSVGSYFHTWVESLVGMYTRGEDVRCAPYDWRRA-----PNEGTFILA-LREKTEE 183
Qy 197 ASNSNNGKPVILLSHSLGLEFVLQNLNPPSWRKKFTKHEIALSAPWGAIDEMYTFAS 256
Db 184 MYQLVGG--PVVLVAHSMGNMTYVFLQRPQAMKDKYIRAVVSLGAPMGVAKTLRYLAS 242
Qy 257 GNTLGVPLVLDLVNDEGRSSESNNMLLPNPKIIGPOKPIYITPRYSADHWDFLKD 316
Db 243 GDMNRIPVIGDKIREQORSVSTSMPLPVNYVTSPEKVFVQTFPTINTLTDYRKFQDI 302
Qy 317 GPFGES--VYFETFRILPLIGNIKAPQVPTICMGCGVGLLEFLYKGKGDGDER--EISYGD 374
Db 303 GFEDGMLRQDTE--GLVEATMPGVQHLCLYGVGVPPDSFYF--ESFPDRDKICRGD 358
Qy 375 GDGTVNLVSLALQSLMKKEKNQYLKVKIDGVSHSTSLKREVAL 419
Db 359 GDGTVNLKSAALQCA--WOSROHOVLQDELPGSEHTEMLNATTL 402

RESULT 3
US-09-978-192A-157
; Sequence 157; Application us/09978192A
; Patent No. US20020177553A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc
APPLICANT: Baton, Dan
APPLICANT: Perrera, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC9
CURRENT APPLICATION NUMBER: US/09/978,192A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079669
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083435
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558

```

; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

```

Query Match 26.1%; Score 610; DB 9; Length 412;

Best Local Similarity 34.8%; Pred. No. 2e-49;

Matches 141; Conservative 80; Mismatches 156; Indels 28; Gaps 12;

```

QY 21 VVVMLSLC--TCGASLMDPLILIPGNGQLARLTNQKPT--FICSWPLIKKN 76
DB 20 LLLLLLMDPALPACRHPVVLVPDGLGQLEAKLD--KPVVYVLS-----KTE 70
QY 77 GWFRLWDSVILAPFTQCAERMTLHYHQLDVFNTPCVETRVVPHFGTSLVLPN 136
DB 71 SYFTILNLELLPVIIIDMIDNIRLVNKTSAQTQPPGCVDRVVGFGTSELEFLDS 130
QY 137 LKIIITVMAPLVDSLOKLGADETLFGAPYDFRYLAAEGHPSQVSKFLDKNLIB 196
DB 131 KSSVGSYFTHTWESLVGWCYTREGEDVRCAPYDWRRA-----PNEGPFYLA-LREMI 183
QY 197 ASNSNNGKPIILSLGSLGFVQLNRPSPSRKKFIKFIKFIKFIKFIKFIKFIKFI 256
DB 184 MWOLYGG-PVVLVAHSMGNMYTLFYQROPQAKWDYIRAFVSLGAPWGVAKTLRVLAS 242
QY 257 GNTLGVPLVPLVDRDESSSESNLMLNPNKIFGQKPIVIRIPYSAHMDVDFLKI 316
DB 243 GDNRRIRPVGPLKIRGOQSAVSTSWLLPYNTWSEKVFVQPTINTYLRDRKRFQDI 302

```

```

QY 317 GFPEG-VVPEYTRILPLIGNIKAPOVPTICMGVGTLETLFYKGDPDFER-PEISYGD 374
DB 303 GFEDGMRLQDTE--GLVEATMPGVLHCLTGTGVTPTPSFY--ESFDRPKICFGD 358
QY 375 GDCIVNLVSLIAQSLMKEKNQYIKVKKIDGVSHSILKDEVAL 419
DB 359 GDTVNLKSALQCOA-WQSRQHOVLQLPGLGSEHTEMLNAVTL 402

```

RESULT 4

US-09-999-832a-157

; Sequence 157, Application US/09999832a

; Publication No. US20020192706A1

; GENERAL INFORMATION:

```

; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Bolstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James.
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC63
; CURRENT FILING DATE: 2001-10-24
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/07450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910

```

;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078939
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079664
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079786
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079920
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/079923
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/080105
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080194
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23

;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 26.1% Score 610; DB 9; Length 412;

Best Local Similarity 34.8%; Pred. No. 2e-49; Indels 28; Gaps 12;

Matches 141; Conservative 80; Mismatches 156;

QY 21 VVWMLSLC--TCGASMLPLILIPGNGNOLEARTLNQKRPST--FICESWYPLIKKN 76
Db 20 LLLMLLADPALPAGRRHPVYLVPGDLGNQLEAKD---KPTVYHYLCS-----KITE 70


```
;; PRIOR FILING DATE: 2001-08-28
;; PRIOR APPLICATION NUMBER: 09/946374
;; PRIOR FILING DATE: 2001-09-04
;; PRIOR APPLICATION NUMBER: PCT/US98/18824
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: PCT/US99/00106
;; PRIOR FILING DATE: 1999-01-05
;; PRIOR APPLICATION NUMBER: PCT/US99/05028
;; PRIOR FILING DATE: 1999-03-08
;; PRIOR APPLICATION NUMBER: PCT/US99/08615
;; PRIOR FILING DATE: 1999-04-20
;; PRIOR APPLICATION NUMBER: PCT/US99/12252
;; PRIOR FILING DATE: 1999-06-02
;; PRIOR APPLICATION NUMBER: PCT/US99/20111
;; PRIOR FILING DATE: 1999-09-01
;; PRIOR APPLICATION NUMBER: PCT/US99/20594
;; PRIOR FILING DATE: 1999-09-08
;; PRIOR APPLICATION NUMBER: PCT/US99/28313
;; PRIOR FILING DATE: 1999-11-30
;; PRIOR APPLICATION NUMBER: PCT/US99/28551
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/28634
;; PRIOR FILING DATE: 1999-12-01
;; PRIOR APPLICATION NUMBER: PCT/US99/30095
;; PRIOR FILING DATE: 1999-12-16
;; PRIOR APPLICATION NUMBER: PCT/US99/30999
;; PRIOR FILING DATE: 1999-12-20
;; PRIOR APPLICATION NUMBER: PCT/US00/00376
;; PRIOR FILING DATE: 2000-01-06
;; PRIOR APPLICATION NUMBER: PCT/US00/03565
;; PRIOR FILING DATE: 2000-02-11
;; PRIOR APPLICATION NUMBER: PCT/US00/04341
;; PRIOR FILING DATE: 2000-02-18
;; PRIOR APPLICATION NUMBER: PCT/US00/04342
;; PRIOR FILING DATE: 2000-02-18
;; PRIOR APPLICATION NUMBER: PCT/US00/05841
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: PCT/US00/06884
;; PRIOR FILING DATE: 2000-03-15
;; PRIOR APPLICATION NUMBER: PCT/US00/08439
;; PRIOR FILING DATE: 2000-03-30
;; PRIOR APPLICATION NUMBER: PCT/US00/13705
;; PRIOR FILING DATE: 2000-05-17
;; PRIOR APPLICATION NUMBER: PCT/US00/14042
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: PCT/US00/14941
;; PRIOR FILING DATE: 2000-05-30
;; PRIOR APPLICATION NUMBER: PCT/US00/15264
;; PRIOR FILING DATE: 2000-06-02
;; PRIOR APPLICATION NUMBER: PCT/US00/22031
;; PRIOR FILING DATE: 2000-08-11
;; PRIOR APPLICATION NUMBER: PCT/US00/23522
;; PRIOR FILING DATE: 2000-08-23
;; PRIOR APPLICATION NUMBER: PCT/US00/30873
;; PRIOR FILING DATE: 2000-11-10
;; PRIOR APPLICATION NUMBER: PCT/US00/32678
;; PRIOR FILING DATE: 2000-12-01
;; PRIOR APPLICATION NUMBER: PCT/US01/06520
;; PRIOR FILING DATE: 2001-02-28
;; PRIOR APPLICATION NUMBER: PCT/US01/06666
;; PRIOR FILING DATE: 2001-03-01
;; PRIOR APPLICATION NUMBER: PCT/US01/17092
;; PRIOR FILING DATE: 2001-05-25
;; PRIOR APPLICATION NUMBER: PCT/US01/17800
;; PRIOR FILING DATE: 2001-06-01
;; PRIOR APPLICATION NUMBER: PCT/US01/19692
;; PRIOR FILING DATE: 2001-06-20
;; PRIOR APPLICATION NUMBER: PCT/US01/21066
;; PRIOR FILING DATE: 2001-06-29
;; PRIOR APPLICATION NUMBER: PCT/US01/21735
;; PRIOR FILING DATE: 2001-07-09
;; PRIOR APPLICATION NUMBER: PCT/US01/27099
;; PRIOR FILING DATE: 2001-08-29
```

```
;; NUMBER OF SEQ ID NOS: 91
;; SEQ ID NO: 8
;; LENGTH: 412
;; TYPE: PRT
;; ORGANISM: Homo Sapien
;; US-10-001-054-8

Query Match
Best Local Similarity 34.8%; Pred. No. 2e-49; Length 412;
Matches 141; Conservative 80; Mismatches 156; Indels 28; Gaps 12;

QY 21 VVVMLESLC--TCGASNDPLILIPNGNGQLEARTLNQKST--FICESWYPLIKRN 76
    ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 20 LILLMLADPALPAGRHPPVLPDGLNQLEAKD---KPTVHYLCS-----KTE 70
    : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 77 GWFRLMFDSSVILAPFTQCFAERMTHLYHOELDDYNTNPGVETRVHPFSSTSLYLNR 136
    : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 71 SYFTIMLNLLELLPVLIIDMIDNIRLVYNTSRATOPPDGVDRVPGFKTSTLEFLDS 130
    : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 137 LKHITGYMAPINDSLKLGADGETLFGAPYDFRYGLAAGHPSQVSKFLDKLIEE 196
    : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 131 KSSVGSYFHTWESLVGWGTYRGEVDVGAAYDWRRA-----PNEGPFYFLA-LREMI 183
    : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 197 ASNSNMGKPYILSHSLGFLVQLNRNPPSWRRKFKIKHFIALSAPMGADIDEMTFAS 256
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 184 MYOLYGG-FVYLVAHSMGNNTYLYFLOROPQAKDKYIRAFVSLGAPMGVAKTLVLAS 242
    : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 257 GNTLGVPLVDPLLVDRDORSSESMLWLPNPKIFGQKPIVYPIRYSAHMDVFLDKI 316
    : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 243 GDNRRIPVIGLPIKIRBOQSAVSTWMLPYNTWSPEKVFQPTINTYLRDRKRFQDI 302
    : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 317 GPPEG-VYPERITPLIGIKAPQVPIICIMGTGVTLETLEFYGGDDER-PELSYD 374
    : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 303 GFEDGWLMDQDE--GLVEXTMPPGVQLCLGTGVPYDPSFY--ESPDRPKICFD 358
    : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 375 GDSVTMLVSLALQSLMKKEKNQYLKVKRIDGVSHTLKDEVAL 419
    ||||| | | | | | | | | | | | | | | | | | | | | | | |
Db 359 GDSVTMLKSLAQCA-WQSRQEHQVLLQELPGEHEIMLANATTL 402
    ||||| | | | | | | | | | | | | | | | | | | | | | | |

RESULT 6
US-09-978-189-157
; Sequence 157, Application US/09978189
; Publication No. US20030004102A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Olang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gertlisen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James:
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
```

TITLE OF INVENTION:	Acids Encoding the Same
FILE REFERENCE:	P2630D1C7
CURRENT APPLICATION NUMBER:	US/09/978,189
PRIOR FILING DATE:	2001-10-15
PRIOR APPLICATION NUMBER:	09/918585
PRIOR FILING DATE:	2001-07-30
PRIOR APPLICATION NUMBER:	60/062250
PRIOR FILING DATE:	1997-10-17
PRIOR APPLICATION NUMBER:	60/064249
PRIOR FILING DATE:	1997-11-03
PRIOR APPLICATION NUMBER:	60/065311
PRIOR FILING DATE:	1997-11-13
PRIOR APPLICATION NUMBER:	60/066364
PRIOR FILING DATE:	1997-11-21
PRIOR APPLICATION NUMBER:	60/077450
PRIOR FILING DATE:	1998-03-10
PRIOR APPLICATION NUMBER:	60/077632
PRIOR FILING DATE:	1998-03-11
PRIOR APPLICATION NUMBER:	60/077641
PRIOR FILING DATE:	1998-03-11
PRIOR APPLICATION NUMBER:	60/077649
PRIOR FILING DATE:	1998-03-11
PRIOR APPLICATION NUMBER:	60/077791
PRIOR FILING DATE:	1998-03-12
PRIOR APPLICATION NUMBER:	60/078004
PRIOR FILING DATE:	1998-03-13
PRIOR APPLICATION NUMBER:	60/078886
PRIOR FILING DATE:	1998-03-20
PRIOR APPLICATION NUMBER:	60/078936
PRIOR FILING DATE:	1998-03-20
PRIOR APPLICATION NUMBER:	60/078910
PRIOR FILING DATE:	1998-03-20
PRIOR APPLICATION NUMBER:	60/078939
PRIOR FILING DATE:	1998-03-20
PRIOR APPLICATION NUMBER:	60/079294
PRIOR FILING DATE:	1998-03-25
PRIOR APPLICATION NUMBER:	60/079656
PRIOR FILING DATE:	1998-03-26
PRIOR APPLICATION NUMBER:	60/079664
PRIOR FILING DATE:	1998-03-27
PRIOR APPLICATION NUMBER:	60/079689
PRIOR FILING DATE:	1998-03-27
PRIOR APPLICATION NUMBER:	60/079663
PRIOR FILING DATE:	1998-03-27
PRIOR APPLICATION NUMBER:	60/079728
PRIOR FILING DATE:	1998-03-27
PRIOR APPLICATION NUMBER:	60/079786
PRIOR FILING DATE:	1998-03-27
PRIOR APPLICATION NUMBER:	60/079920
PRIOR FILING DATE:	1998-03-30
PRIOR APPLICATION NUMBER:	60/079923
PRIOR FILING DATE:	1998-03-30
PRIOR APPLICATION NUMBER:	60/080105
PRIOR FILING DATE:	1998-03-31
PRIOR APPLICATION NUMBER:	60/080107
PRIOR FILING DATE:	1998-03-31
PRIOR APPLICATION NUMBER:	60/080165
PRIOR FILING DATE:	1998-03-31
PRIOR APPLICATION NUMBER:	60/080194
PRIOR FILING DATE:	1998-03-31
PRIOR APPLICATION NUMBER:	60/080327
PRIOR FILING DATE:	1998-04-01
PRIOR APPLICATION NUMBER:	60/080328
PRIOR FILING DATE:	1998-04-01
PRIOR APPLICATION NUMBER:	60/080333
PRIOR FILING DATE:	1998-04-01
PRIOR APPLICATION NUMBER:	60/080334
PRIOR FILING DATE:	1998-04-01
PRIOR APPLICATION NUMBER:	60/081070
PRIOR FILING DATE:	1998-04-08
PRIOR APPLICATION NUMBER:	60/081049
PRIOR FILING DATE:	1998-04-08
PRIOR APPLICATION NUMBER:	60/081071
PRIOR FILING DATE:	1998-04-08

[illegible]


```

RESULT 7
US-09-919-497-80
; Sequence 80, Application US/09919497
; Patent No. US2002010662A1
; GENERAL INFORMATION:
; APPLICANT: Muller, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/77225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 80
;
LENGTH: 440

```

```

RESULT      B
US-09-738-626-3594
# Sequence 3594, Application US/09738626
# Publication No. US20020197605A1
# GENERAL INFORMATION:
# APPLICANT: NAKAGAWA, SATOSHI
# APPLICANT: MIZOGUCHI, HIROSHI
# APPLICANT: ANDO, SEIKO
# APPLICANT: HAYASHI, MUKIRO
# APPLICANT: OCHIAI, KEIKO
# APPLICANT: YOKOI, HARUHIKO
# APPLICANT: YATEISHI, NAOKO
# APPLICANT: SENOH, AKIHIRO
# APPLICANT: IKEDA, MASATO
# APPLICANT: OZAKI, AKIO
# TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
# FILE REFERENCE: 249-125
# CURRENT APPLICATION NUMBER: US/09/738, 626
# CURRENT FILING DATE: 2000-12-18
# PRIOR APPLICATION NUMBER: JP 99/377484
# PRIOR FILING DATE: 1999-12-16
# PRIOR APPLICATION NUMBER: JP 00/159162
# PRIOR FILING DATE: 2000-04-07
# PRIOR APPLICATION NUMBER: JP 00/280988
# PRIOR FILING DATE: 2000-08-03
# NUMBER OF SEQ ID NOS: 7059
# SOFTWARE: PatentIn ver. 3.0
# SEQ ID NO 3594
# LENGTH: 300
# TYPE: PRT
# ORGANISM: Corynebacterium glutamicum
# US-09-738-626-3594

```

Query Match 4.8%; Score 112; DB 9; Length 300;
Best Local Similarity 20.1%; Pred. No. 0.0085;
Matches 76; Conservative 53; Mismatches 107; Indels 142; Gaps 21;

```
QY 96 FAERMTLHYHOELDDYFNTPGVETRVPHFGSTNSLLYLNPRLKHITGYMA----- 145
||| | ||| | : ||| | ||| | :
Db 28 FAERT-----QELVDH-----ATNITLPLTALKP-KGFFQODMRARPSGA 68
QY 146 ---PLV-----DSLQKLG---YADGETLFGAPYDFRYGLAEGHPSOVGSKFLKDL 190
||| | ||| | : ||| | ||| | :
Db 69 RPPVIVLHSGASGASKEWEGAEIERSKGMVFAFDEFTGRATEPIAASAOIGAYI--DA 126
QY 191 KNLLEASNSNNGKPVILLSHSISGLFVLQLLNNRPPSMRKKF-----IKHPIALASAP-- 243
||| | ||| | : ||| | ||| | :
Db 127 VLVYGAAGQ-----YLVGHSGGVARY-----WMRTYGGYMKVRHMISTEPNH 172
QY 244 ---MGAIIDENYTRASGNTLGPVLDPLVDEORSSESNLMLFPKIFGPOKEVITP 300
||| | ||| | : ||| | ||| | :
Db 173 GTLGGILNPMTKVKSSE---GTIEKLMNR-----LFGP-----TG 205
QY 301 IRPISAHDMVPLKDIGPPEGVYPERIKILPLIGNIKAPQVPTICMGTVGTLLETLYG 360
||| | ||| | : ||| | ||| | :
Db 206 FEOLRGHDIIEFLAD-----GDDLDPGVYTCI-GTHF----- 237
QY 361 KGDEPEREISY--GDGDGVNLVSLALQSLMKKEKNQYLVKVKYIDGVSHTSILKDEVA 418
||| | ||| | : ||| | ||| | :
Db 238 -DPTIQPEVAFLEVNEDEDDPRRY-----HNIVWEDEHPRAM-----IAHNDVREDRV 285
QY 419 LNEIVGEITSINSHAEIG 436
||| | ||| | : ||| | ||| | :
Db 286 IEIVRAELDRY---ARIG 300

RESULT 9
US-09-950-368-17
; Sequence 17, Application US/09950368
; Patent No. US20020061580A1
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
; FILE REFERENCE: GC511-PCR
; CURRENT APPLICATION NUMBER: US/09/950,368
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 09/355,166
; PRIOR FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Bacillus
US-09-950-368-17

Query Match 4.6%; Score 106.5; DB 10; Length 314;
Best Local Similarity 21.9%; Pred. No. 0.03; Indels 89; Gaps 12;
Matches 65; Conservative 34; Mismatches 109; Indels 89; Gaps 12;

QY 71 LIRKKNCFRLMFDSSVILAFETOCFAERMVLYHQLDDYFNTPGVETRVPH----- 123
||| | ||| | : ||| | ||| | :
Db 4 LVDSENHHTLRFs---LRGMSYCMKEQTDRNGCTSNMFTJPGTEVRMSSNENRT 60
QY 124 -----FGSTNSLLYLNPRLKHITGYMAPLVDSLQKLG 155
||| | ||| | : ||| | ||| | :
Db 61 YHIFISKRPPRPAGYRVIVLIDANSVFGTMTAVRIGGRRPKEGTGVIPIAVIG--IG 117
QY 156 YADETELEGARY-DPRYGLAEGHPSOVS-----KFL-DLNLLEEASNS 200
||| | ||| | : ||| | ||| | :
Db 118 YETAEPPSSAHRDPTMTAOSKLPERPDRGEMPBHGAECGFPRTEEDLPELERYOI 177
QY 201 NNGKRVILLSHSLGCLFVLQLLNNP-----PS--WRKKFI-----KHFIALSAPW 244
||| | ||| | : ||| | ||| | :
Db 178 DKRQGTI-FGHSISGLFVLQVLLKRPDAFYIAGSFSIHNNKFFILKKTHTHEFVSLTKN 236
QY 245 CGAID-----EMVTFASNTGLVPLVDPLVDEQR-----SSESNLMLLP 285
||| | ||| | : ||| | ||| | :
Db 237 NOPINILLAGELBQHHSKRMNDNARELYERLAVLSEQIGRAERCFEGEGHISVLP 293
```

```
RESULT 10
US-09-738-626-5970
; Sequence 5970, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIJO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIJO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5970
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5970

Query Match 4.4%; Score 103.5; DB 9; Length 338;
Best Local Similarity 20.7%; Pred. No. 0.063;
Matches 72; Conservative 47; Mismatches 119; Indels 109; Gaps 16;

QY 135 PRKHITGYMAPLYDS-----LQKIGYADGETLFGAPYDRRYG--LAAEGHPSOVS--S 184
||| | ||| | : ||| | ||| | :
Db 41 PGLKHVDRH--TIVDSGLSHITYVNGHAENAT--ATVFIHGFTLLAAEVYVQVBYLQ 95
QY 185 KFLDKLN-LIEEASNSNNGK-----PVILLSHSLGCLF 217
||| | ||| | : ||| | ||| | :
Db 96 TFFYNIKSVILIDARGHGATGQIRELCITECTANDVLAHNEHAPYTGPLLIVGHSLSGLT 155
QY 218 VLQLLNNRPPSMRKKFIKHFIALSAPWCGAIDENYTRASGNTLGPVLDPL-----LVR 271
||| | ||| | : ||| | ||| | :
Db 156 ALNLVKKRADHSLRKRIVGMVLVAT-----STESISTGCLPQVLAAPLADNIKNAVEAAPN 210
QY 272 DEQRSSENLMLLNPKIFGPOKPIVITPIRPSAHD-MVDFLAD-----IGPPE 320
||| | ||| | : ||| | ||| | :
Db 211 DAQFRQYATTFL-----APTLATAVFORDTNDVIDFHAAHMHETPLDFVGGFED 261
QY 321 GVPPY-ETRIPLIGNIKAPQVPTICMGTVGTLLETLYGKGDPDEREISYDGDGTV 379
||| | ||| | : ||| | ||| | :
Db 262 DLQEHDELDAAPALEGLKG-----YVLAGELDVTPTISQAD----- 297
QY 380 NLVSLALQSLMKKEKNQYLVKVIDGVSHTSIILKDEVALNEIVGEI 426
||| | ||| | : ||| | ||| | :
Db 298 -----RICEVWPGARLOI-----AEGAGHMLPLEARGLINNAIGNI 333

RESULT 11
US-09-748-739A-22
; Sequence 22, Application US/09748739A
; Patent No. US20020119489A1
; GENERAL INFORMATION:
; APPLICANT: Lockridge, Oksana
; APPLICANT: Watkins, Jeffrey D.
; TITLE OF INVENTION: Butyrylcholinesterase Variants and
; Methods of Use
```

```

; FILE REFERENCE: P-IX 4143
; CURRENT APPLICATION NUMBER: US/09/748,739A
; CURRENT FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 574
; TYPE: PRF
; ORGANISM: Felis catus
; US-09-748-739A-22

```

Query Match	4.4%	Score 101.5;	DB 10;	Length 574;
Best Local Similarity	22.7%	Pred. No. 0.21;		
Matches 99;	Conservative 56;	Mismatches 150;	Indels 131;	Gaps 27

[illegible]

Query Match	4.38;	Score 100.5;	DB 10;	Length 342;
Best Local Similarity	20.68;	Pred. No. 0.12;		
Matches 85;	Conservative 36;	Mismatches 156;	Indels 135;	Gaps 18

QY	71	LIKKNGEFLMPDS-----SVIAPPTOCFAEBMFLNH-----	HOE	107
Db	5	LEOOSOGWLSWLEPTMRPTSOLKKNYEARILLOCLQNKFLARIVSLPNOKIMTWTYSPE	64	
QY	108	LDDEYFNTGVSTRPHFGSTNSLTYLNPRLKHTGYMAPLVDSLOKGAYADGETFLGAPY	167	
Db	65	QND--RRPLV--VWHGEGGGVWILN-----MDLSARRTLHTFDLIGFGR	107	
QY	168	DPRYGLAEGHPSOVGSFKELDKNLIEBANSNNGKP-VILLSHSLGGLFVLQLLRNP	226	
Db	108	SSRAFPED--PEGADEFEVTSI-----ETWRETMGIPSMILGSHSLGGLATSYSIKYP	160	
QY	227	PSMRKFKIHNIALSAPMGAI-----DEMYFASGNTLGVLPVLYRDQRS	276	
Db	161	DR-----VKHIL-LVDPGEPFLRPNSEIRAPPAWKAASVJG--RSNPFLAVL----	207	
QY	277	SESNTMLLPNKIRGPOKPIYITREPSADHWDFLDIGFEGVYPTFR-----	328	
Db	208	-----RVAGPMGPGVQGFRRDFPKKFADEFEDDTISYIYTHCAQNPPSGTAF	256	
QY	329	--ILPLIGNIKAP-----OVPTICJMG--TGVSLETLEFYGKGFDERPEISYGD	374	
Db	257	KAMSEFEGMARPMLEIRHLIRKQVPIPTMIYGSPTWIDTST---GKVVAKORPD-----	307	
QY	375	GDGYVNLVSLALDOSLMEKNQVLKTVKIDGVSHTSILDEVALNRYGEI	426	
Db	308	-----SYVRMEIKGASHVAYADOPHLENNVVEI	337	

LENGTH: 346 AMINO ACIDS
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-027-805-34

Query Match 4.2% Score 97.5; DB 9; Length 346;
Best Local Similarity 25.1%; Pred. No. 0.24;
Matches 42; Conservative 24; Mismatches 64; Indels 37; Gaps 8;

QY 143 YAAPLVDSLOKGYADGETLFGAPYDFRGLAEGHPSQVSKFLKDKNLLEASNSNN 202
DB 142 YAAVLARELVEWGYP-----VVVFDFR-GHGESSGSTTIGPREVIDARAAYGVYSERFP 194
QY 203 GKPVLLSHSLGGLFVL-----QLNRPNSMR-KKFKHFTALSAP---MGCATD 249
DB 195 GKRILLVGSMSGANAIVAGADDPRIYAAADSPYRLRDVLRPMLEKTPPLPGWGVLA 254
QY 250 EMY-TFASGNTLGV-----PLV-----DPLLVRDEQRSSSES 279
DB 255 GFYGRMLAGVDLGFSPAGYGRVDKPLLVYVYGRDPLVTDEARSLAS 301

RESULT 14
US-09-903-410-34
Sequence 34, Application US/09903410
Patent No. US20020146799A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: ROBERTSON, Dan
APPLICANT: MURPHY, Dennis
APPLICANT: REID, John
APPLICANT: MAFFIA, Anthony
APPLICANT: LINK, Steven
APPLICANT: SWANSON, Ronald
APPLICANT: WARREN, Patrick
APPLICANT: KOSMOTKA, Anna
TITLE OF INVENTION: ENZYMES HAVING ESTERASE ACTIVITY AND METHODS OF USE THEREOF
FILE REFERENCE: DIVER1180-2
CURRENT APPLICATION NUMBER: US/09/903,410
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 09/382,242
PRIOR FILING DATE: 1999-08-24
PRIOR APPLICATION NUMBER: US 08/602,359
PRIOR FILING DATE: 1996-02-16
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn version 3.0
SEQ ID NO 34
LENGTH: 346
TYPE: PRT
ORGANISM: Pyrodicticum
US-09-903-410-34

Query Match 4.2% Score 97.5; DB 10; Length 346;
Best Local Similarity 25.1%; Pred. No. 0.24;
Matches 42; Conservative 24; Mismatches 64; Indels 37; Gaps 8;

QY 143 YAAPLVDSLOKGYADGETLFGAPYDFRGLAEGHPSQVSKFLKDKNLLEASNSNN 202
DB 142 YAAVLARELVEWGYP-----VVVFDFR-GHGESSGSTTIGPREVIDARAAYGVYSERFP 194
QY 203 GKPVLLSHSLGGLFVL-----QLNRPNSMR-KKFKHFTALSAP---MGCATD 249
DB 195 GKRILLVGSMSGANAIVAGADDPRIYAAADSPYRLRDVLRPMLEKTPPLPGWGVLA 254
QY 250 EMY-TFASGNTLGV-----PLV-----DPLLVRDEQRSSSES 279
DB 255 GFYGRMLAGVDLGFSPAGYGRVDKPLLVYVYGRDPLVTDEARSLAS 301

RESULT 15
US-09-738-626-4656

Sequence 4656, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 4656
LENGTH: 1014
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-4656

Query Match 4.2% Score 97.5; DB 9; Length 1014;
Best Local Similarity 22.9%; Pred. No. 1.1;
Matches 79; Conservative 36; Mismatches 109; Indels 121; Gaps 19;

QY 150 SLQKLGADGETLFGAPYDFR-----YGLAEGHPSQVSKFLK--DLKNLIEASN 199
DB 370 SMPKIQNIDGE-----PMDEKESPMSMALRGQGVGLAKLGEVLGEDPDLARRIFEISA 424
QY 200 SNNKRPVLLSHSLGGLFVLQNLNPNPMSRRKKFKHFTALSAPGCAIDEMYT----- 253
DB 425 S-----PMYLRGETEPG-----HALVT---WHDSTNEYTMOQLTL 457
QY 254 -FASGNTL-----GVPLVDPPLVRDEQRSSSESNLMLPNPKIFGPQRIYTPPIRPSA 306
DB 458 AYEESRLFEKAPQGIAMD-----SGE-----IWMANSFDDL--VGTTPVRLRGL 503
QY 307 HDVDFLKDIFPEGVYPERIL-----PLIGNIRAPQVPTICMGTVG 352
DB 504 N-----LEDGVEEGTMEYVTEVLSDPEAVVHLDRSLFTLRGKQKNVAMSFSS--MGNVG 557
QY 353 TLETLFYGKGDPERPEI-----SYGDGCVTVNLVSLALOS-----LWKKEKN----- 396
DB 558 RIGTLVNVVDYTERQELIEVEHLADHDSLTGLVNRRLLEDIDIELLIKNERDSTDSAL 617
QY 397 -----QYLKVKIKDGVSH-----TSILKDEVALNEIVEGI 426
DB 618 LLDDIDYFEVN-DSLGHAGDQLLIEFAELIKDSVROSDIVGRI 661

Search completed: February 11, 2003, 19:20:22
Job time : 32 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 19:13:15 ; Search time 21 Seconds
(without alignments)
2027.979 Million cell updates/sec

Title: US-09-857-612A-14

Perfect score: 2333
Sequence: 1 MKKEQEGSLKEVALTVTV.....GEITSINSHALGSLNLESG 443

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR-73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	541	23.2	440	1 XXHUN	phosphatidylcholin
2	537	23.0	440	2 JC1502	phosphatidylcholin
3	516	22.1	438	1 XXMSN	phosphatidylcholin
4	515	22.1	413	2 I50662	lecithin-cholester
5	515	22.1	440	1 XXRTN	phosphatidylcholin
6	512.5	22.0	417	2 T23726	hypothetical prote
7	457.5	19.6	272	2 T41755	hypothetical prote
8	181	7.8	623	2 T40685	phosphatidylcholin
9	163	7.0	655	2 T51780	hypothetical prote
10	147	6.3	562	2 C86171	hypothetical prote
11	125.5	5.4	493	2 T04767	hypothetical prote
12	113	4.8	600	2 E86854	hypothetical prote
13	109	4.7	984	2 T44608	pyruvate carboxyla
14	106.5	4.6	314	2 B70013	hypothetical prote
15	106	4.5	586	1 S73355	exonuclease ABC c
16	105.5	4.5	529	2 AD1042	hypothetical prote
17	105	4.5	757	2 T38124	hypothetical prote
18	104	4.5	317	2 T00550	probable phosphol
19	102	4.4	249	2 H87675	hydroxylase, alpha/b
20	102	4.4	542	2 JN0438	carboxylesterase (
21	101.5	4.4	1407	2 T28702	probable polyketid
22	101	4.3	771	2 T47071	hypothetical prote
23	99.5	4.3	434	2 AD0014	hypothetical prote
24	99.5	4.3	732	2 AD0014	primosomal protein
25	99	4.2	605	2 R35216	FRP6 protein - fow
26	98.5	4.2	830	2 T07824	1,4-alpha-glucan b
27	98.5	4.2	1626	2 T09271	probable tail comp
28	98	4.2	989	2 S69711	hypothetical prote
29	97.5	4.2	390	2 T48524	lysophospholipase-

30	97.5	4.2	861	1 S34730	1,4-alpha-glucan b
31	96.5	4.1	331	1 A70032	conserved hypothet
32	96	4.1	631	2 AH0317	penicillin-binding
33	95.5	4.1	286	2 D75217	probable 2-acetyl-
34	95.5	4.1	3169	2 T00296	toxin B - Escheric
35	95	4.1	297	2 C83311	hypothetical prote
36	95	4.1	357	2 D69820	hypothetical prote
37	95	4.1	446	2 H97091	chitinase family p
38	94.5	4.1	476	2 F81340	glutamate-amonnia
39	94.5	4.1	1137	2 E86708	pyruvate carboxyla
40	94	4.0	462	2 T01732	urp-glucose glucos
41	94	4.0	1019	2 C96519	probable disease r
42	93.5	4.0	209	2 B82501	hypothetical prote
43	93.5	4.0	1269	2 A90267	proteinase related
44	93	4.0	373	2 F72602	probable acetylpol
45	93	4.0	417	1 TVCRGC	phosphoglycerate k

ALIGNMENTS

RESULT 1
XXHUN
phosphatidylcholine-sterol O-acetyltransferase (EC 2.3.1.43) precursor [validated] - hu
N:Alternate names: lecithin-cholesterol acyltransferase precursor; phospholipid-chole
C:Species: Homo sapiens (man)
C:Date: 04-Dec-1986 #sequence-revision 04-Dec-1986 #text-change 03-Jun-2002
C:Accession: A00571; A25575; A29661; J00036; A29133; I52260; A28511
R:McLean, J.; Fielding, C.; Drayna, D.; Dieplinger, H.; Baer, B.; Kohr, W.; Henzel, W
Proc. Natl. Acad. Sci. U.S.A. 83, 2335-2339, 1986
A>Title: Cloning and expression of human lecithin-cholesterol acyltransferase cDNA.
A:Reference number: A00571; MUID:86205550; PMID:3458198
A:Accession: A00571
A:Molecule type: mRNA
A:Residues: 1-440 <MCCL1>
A:Cross-references: GB:M12625; NID:9187022; PIDN:AAA59498.1; PID:9307117
R:McLean, J.; Wion, K.; Drayna, D.; Fielding, C.; Lawn, R.
Nucleic Acids Res. 14, 9397-9406, 1986
A>Title: Human lecithin-cholesterol acyltransferase gene: complete gene sequence and
A:Reference number: A25575; MUID:87091568; PMID:3797244
A:Accession: A25575
A:Molecule type: DNA
A:Residues: 1-440 <MCCL2>
A:Cross-references: GB:X04981; NID:934286; PIDN:CMA28651.1; PID:934287
R:Rogne, S.; Skretting, G.; Larsen, F.; Myklebost, O.; Mewar, B.; Carlson, L.A.; Holm
Biochem. Biophys. Res. Commun. 148, 161-169, 1987
A>Title: The isolation and characterisation of a cDNA clone for human lecithin:choles
A:Reference number: A29661; MUID:88049652; PMID:2823801
A:Accession: A29661
A:Molecule type: mRNA
A:Residues: 13-440 <ROG>
A:Cross-references: GB:M17959; NID:9187026; PIDN:AAA59500.1; PID:9386858
R:Tata, F.; Chaves, M.E.; Markham, A.F.; Scrace, G.D.; Waterfield, M.D.; McIntyre, N.
Biochem. Biophys. Acta 910, 142-148, 1987
A>Title: The isolation and characterisation of cDNA and genomic clones for human leci
A:Reference number: A80666; MUID:88050546; PMID:2823898
A:Accession: J00036
A:Molecule type: mRNA
A:Residues: 17-256, 'H', 258-440 <TAT>
A:Cross-references: GB:X06537; NID:934284; GB:M26268; NID:9187024; PIDN:AAA59499.1; P
A>Note: the authors translated the codon CAT for residue 241 as ile and CAG for resid
R:Yang, C.; Manoochian, D.; Pao, O.; Lee, F.; Knapp, R.D.; Gotto Jr., A.M.; Pomnall, H
J. Biol. Chem. 262, 3086-3091, 1987
A>Title: Lecithin: cholesterol acyltransferase. Functional regions and a structural m
A:Reference number: A29133; MUID:87137578; PMID:2880847
A:Accession: A29133
A:Molecule type: protein
A:Residues: 25-284, 'Q', 286-333, 'Q', 335-440 <YAN>
R:Buyo, H.; Kusunoki, J.; Ogasawara, M.; Yamamoto, T.; Ohta, Y.; Shimada, T.; Satou,
Biochem. Biophys. Res. Commun. 181, 933-940, 1991
A>Title: Molecular defect in familial lecithin:cholesterol acyltransferase (LCAT) def
A:Reference number: I52260; MUID:92109783; PMID:1662503
A:Accession: I52260


```
Oy 393 EEKNQYLKVVKIDGVSHTSILKDEVALNEI 4222
      ::| : ::::| | ::|
Db 385 GRQSAVHLRLRMNGTDHLMVFSNKTLEHI 4144
```

RESULT 6
T23726
hypochemical protein M05B5.4 - *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T23726
R:Gardner, A.

submitted to the EMBL Data Library, April 1996
A: Reference number: Z19790
A: Accession: T23726
A: Status: preliminary; translated from GB/EMBL/DDJ
A: Molecule type: DNA
A: Residues: 1-417 <Wt>
A: Cross-references: EMBL:Z19726; PIDN:CAA9583.1; GSPDB:GN00019; CESP:M05B5.4
A: Experimental source: Clone M05B5

A: Introns: 39/1; 92/2; 165/1; 212/3; 262/2
C: Superfamily: phosphatidylcholine-sterol acyltransferase

Query Match	22.0%	Score 512.5;	DB 2;	Length 417;
Best Local Similarity	32.8%	Pred. 0.3.8e-33;		
Matches 129; Conservative 79; Mismatches 160;			Indels 25;	Gaps 12;

[illegible]

```

0y 156 YADGETLFGAPYDFRYGLAEGHPSQVGSKFLDKLKNLIEASNSNNGKPVILLSHSLGG 215
    | : : : : : | : : : : | : : : : | : : : : |
Db 147 YRRKKNVIGAPFEDMKRS-----PNEL-NDYLLQLKSLLETTYRRMWDNQKILVGHSMGN 199

```

Qy 216 LEVLIQLLNRR-PPSMRRKEIKHFEALFALSPPMGALIDEMETFGASGTLVQ--PLVDFLLVYD 272
 200 PLSLYLENNYDQAMKDKYISIFVSLAAPMAGSMQIYRLRFSQYNNMYNRYRLILPPSSLRA 255
 Db 273 EQRSSSENLMMLNPKPIFGPOKPYVYPIRYSASHDVAVDLAKDIGFEGCVYPIYETRIPLP 333
 260 MORFSSSALEPSPVAMKPEHILATYADKKYNTVGNIKKEFQDIDNYNVMGEOYX-QAARL 318
 Qy 216 LEVLIQLLNRR-PPSMRRKEIKHFEALFALSPPMGALIDEMETFGASGTLVQ--PLVDFLLVYD 272
 200 PLSLYLENNYDQAMKDKYISIFVSLAAPMAGSMQIYRLRFSQYNNMYNRYRLILPPSSLRA 255
 Db 273 EQRSSSENLMMLNPKPIFGPOKPYVYPIRYSASHDVAVDLAKDIGFEGCVYPIYETRIPLP 333
 260 MORFSSSALEPSPVAMKPEHILATYADKKYNTVGNIKKEFQDIDNYNVMGEOYX-QAARL 318

OY 333 IGNKAAPOVPTTCMGNGVGCTLETLFPGKGDFDRPREST-Y-GDSDGYNVNLVSLLALOSLM 391
| : | : | : | : | : | : | : | : | : | : | : | : | :
Db 319 NGNLSGQVPHCHCYGTGGVTPPEAKFSMAPEGFPEDYPTEFGGDODGRYNKKSATVCIN-W 377

OY 392 KEENKOY-LKYVKIDGVSHSLDLDEVALNEIV 423
| : | : | : | : | : | : | : | : | : | : | : | : | :
Db 378 IGNNNGIKVTLYHEVFQADHMALKHPNAI-ELV 409

RESULT 7
T14755
hypothetical protein DKFP564A0122.1 - human
C:Species: Homo sapiens (man)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999
C:Accession: T14755

A:Reference number: Z18181
A:Accession: T14755
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-272 <NAME>
A:Cross-references: EMBL:AL110209

•

RESULT 7
T14755
hypothetical protein DKFZp564A0122.1 - human

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14755
R:Wambutt, R.; Heuvelink, D.; Mewes, H.W.; Gassenhuber, J.; Wilmann, S.
submitted to the Protein Sequence Database, August 1999

A:Experimental source: fetal brain; clone DKFzp564A01222
C:Genetics:
A:Note: DKFzp564A0122.1

Query Match	19.68	Score 457.5	DB 2	Length 272
Best Local Similarity	37.18	Pred. No. 5.2e-29		
Matches 102; Conservative	54	Mismatches 104	Indels 15	Gaps 8

```

Oy 147 LVDSLOKLGVADGETLFGAPYDERYGLAEHGPSVGSKLTKDLKLIIEASNSNMGKPV 206
      ::::  ::  ::  :::::::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db 1 MVESLVGWGIVRGEDVGAPYDMKRA-----PENGPPFLA-LRMIIEEMQLYGC-pv 52

```

QY 207 ILLSHSLGLEFVQLLRNPPSRKKFIKHFIALSAPMGCAIDEMYTFASGNTLGVPLVD 2666

Db 53 VLVANSMGMNMTLYFLQROPOAKDKYIRAFVSLGAPMGVAKTLRYLASGNNRIPYIG 1122

QY 267 PLVAVDEQRSSSESNLWLLPNPKIFGQKPIVITPIRPSYASHDMVDCLKDIGFPEG-VYPY 329

Db 113 PLKIRFDQRSVSTSWLLPYNYTWSPKEYEVOPTINVTLRDYYRRFFODIGEDGWLMRO 172

QY 326 ETRILPLIGNIKAPQVPTICIMGTGVLTETLFYGGKDFPER-PEISYGDDGDTVNLSL 384
 : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 173 DTE--GLVEATMPPGVOLHCLVGSTGVPDPDSFYV--ESFPDRDPKCFGGDGSTVNRSA 228

QY 385 LALQSLWKEFNQYLKVKIKDGYSHTSILKDEVAL 419
| | : | : : : : | : | : |
Db 229 LQCOA-WOSRREHOVLLOELPGSEHTEMIANATTTI 262

RESULT 8

phosphatidylcholine-sterol acyltransferase homolog - fission yeast (Schizosaccharomyces pombe)
C.Species: Schizosaccharomyces pombe
C.Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 07-Dec-1999
C.Accession: T40665
R.Layne, M.; Rajandream, M.A.; Barrell, B.G.; Wedler, H.; Kitzner, M.; Mambutti, R.
Submitted to the EMBL Data Library, January 1999
A.Accession: T40685
A.Reference number: Z21944

A: Molecule type: DNA
A: Residues: 1-623 <LVN>
A: Cross-references: EMBL:AL035263; PIDN:CAA22887.1; GSPDB:GN00067; SPDB:SPBCT76.14
A: Experimental source: strain 972h; cosmid c76
C: Geneticks
A: Gene: SPDB:SPBCT76.14
A: Map position: 2

```

Query Match      7.88;   Score 181;   DB 2;   Length 623;
Best Local Similarity 22.1%;   Pred. No. 2,3e-06;
Matches 109;   Conservative 65;   Mismatches 171;   Indels 148;   Gaps
0y 38 P L I I P G G G Q L A R L T N Q K P S T F I C E S W ----- Y P L I K K K N G F R L M P D S S V I L A P 91
| : : : : |

```

DB 147 FVHMFQVJSSOE L5MSENNCSLIFNN ALMGSM3MLNAM 100
QY 92 F--TQCFAEKMTLHYHQELDDYFNTPGVETRVPH-FGSNSLLYLNPRLKHITGY--MAP 146

0Y LVDSIQKLGAVDGETLFGAPYDERFYGLAEDBHQVSQVSKELDKLN.IEESNSNNCKPV 206
147 ::::: ||| : | ||| - - - : | | | | | | | |
Db VLENLAAGT-EPNNMNASIDYMLSTA----NLEEDKTFSTIKMTI-EYSNVHKKV 207
0Y ILLSHSLGLEGVLQIL-----NRNPSSRKKFIKHFTALSPMGAGIDEMVTFAAG- 257
: ||||: | | : | | | | | | |

QY 258 -NTLGVPLVDPLLVRDEQRSSSNL-----WLLPNKIFGPQ----- 293
::: ::| :| :| :|

250 MANUALLY IDENTIFIED SIGNALS IN THE SIGNAL IN 407

QY 294 KPIVITPIRPYSAHMDVFLKDI-----GFPEGVYPTETRL-----P 331
DB 408 RRDIDKDHDFIDDALQFLKAVTFDDDEKVMYLAKEVLSHLATKEKVLKNNMPSKWIWP 467
QY 332 LIGNIK-APQVPITCIMGTVGTFLETFYKGDFFDERPE-----ISY 372
DB 468 LETSIYPAPDMKICVYHGKVPTEGYY-----TNNPEGQPVIDSVDGTFKVENGIYW 522
QY 373 GGDGDTVMNLVSL-LALQSLMKREKNQYLVKVIDGVSHTSI-----LKDEVALNEIYV----- 424
DB 523 DGDGDTPLPLALGLVCNKWKVQ-----TKRENPAINTSTNYETIKHEPPAFLDGLGPR 573
QY 425 --EITSINSHAEI 435
DB 574 SAEHVDILGHSEL 586

RESULT 9

hypothetical protein F28D10_20 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C:Accession: T51780
R:Deiseny, M.; Berger, C.; Cooke, R.; Grellet, F.; Laudie, M.; Mewes, H.W.; Rudd, S.; Le
submitted to the Protein Sequence Database, August 2000
A:Reference number: 225454
A:Accession: T51780
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-665
A:Cross-references: EMBL:AL391254
A:Experimental source: cultivar Columbia; BAC clone F28D10
C:Genetics:
A:Map position: 3
A:Insertions: 132/2; 206/3; 302/2; 415/3; 504/2
A>Note: F28D10_20

Query Match

Best Local Similarity 7.0%; Score 163; DB 2; Length 665;
Pred. Local Similarity 19.0%; Pred. No. 6.9e-05;
Matches 117; Conservative 81; Mismatches 164; Indels 254; Gaps 28;

QY 32 GASNDPLILPG--NGGNQL-EARLTNQKPSFICESWYPLIKKNGMFRIM--FDS 86
DB 86 GYKAFHPVILVPGIYTGLEMEGRPCAE-----GLFRK-----RLMGASFSE 128
QY 87 VLAFTQCFARMTLHYHQELDDYFNTPGVETR-VPHGSGTNSLLYLPRLKHITGYMA 145
DB 129 ILRRLP-CWLEHLSIDSETGLD---PSGIVRAVPPGLVAD-----YPA 168
QY 146 P-----LVDSLQKLVYADGETLFGAPYDFRYGLAEGHPQOVGSKFLDKNLIEAS 198
DB 169 PCYFAMAVLITELAKIGY-EGKNLMASIDWRLSF---HNTEVRDQSLRLKSKTELWY 223
QY 199 NSNNGKPVILLSHSLGGLFVLIQNLN--RNP-----PSWRKKEIKHFIALSAPWG- 246
DB 224 AINGFKKVVVYHSGMGAIFYELHFLKMWETPLPDGGGGGPGCAKHIKSVNINGPALFV 283
QY 247 --AIDEMT-----FASGNTLGVPLVPLVR----- 271
DB 284 PKAVSNLSAEGKDLAYANSLAPGLDSELLKQTLLEHLMRSHSWSIVSLPKGGEAI 343
QY 272 -----DEORSE-----SNLMLLPMPKTEGPOK----- 294
DB 344 WGDLSHAEEGLNCIYSKKSSQLSLNHLKONYSILKPSRKVEPKYGRIVSPGKRASE 403
QY 295 -----PIVITPIRPYSAHMDVFL 313
DB 404 LPSSQLSTLVNKEISRVGDSNDSTSCGEFWESEYEMSNRESITVKAENATATATVLDLL 463
QY 314 KDIG-----FPEGV-----YVETRLPLIGNIKAPQVPITCIM 347
DB 464 RPIARMKMRRAEAFHSHGIADLDLDPKYGHYKYWNSPLETK-LP-----EAPMEMEYCLY 517

QY 348 GTGVGTLETFLFY-----GKGDPE-----RPEISYGDGDTVMNLVSL-- 384
DB 518 GVGIPIPERYIYIKLATSSCKCKSSIPFRIDGSLDGDVCLKGSTRADDDESYPVISAQF 577
QY 385 -----LALQSLMKREKNQ-----YLKVKYKIDGVSHTSLKDEVALNEIYGEI 426
DB 578 MCAKMGKGRKTRFNPSCMDQFLEBYKHKHPGSLLESNGTSGSAHVDM-GNVGLIEDVRLI 636
QY 427 TTSINSHAEIGLSNIFS 442
DB 637 AAGASQGEIGDRITYS 652

RESULT 10

hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C86171
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: C86171
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-352 <STO>
A:Cross-references: GB:AE005172; NID:q4204287; PIDN:AAD10668.1; GSPDB:GNO0141
C:Genetics:
A:Map position: 1

Query Match

Best Local Similarity 6.3%; Score 147; DB 2; Length 552;
Pred. Local Similarity 24.1%; Pred. No. 0.001;
Matches 75; Conservative 48; Mismatches 124; Indels 64; Gaps 16;

QY 8 GLKIEVATLVVVVYVYVMSLCTCGA-----SNLDPLILPGNGNQLA- 52
DB 2 GANSKSVTASFYIAVFLI---CGGRVAVDETEHGGYKSLSG-IITIPGASTYDLRW 57
QY 53 -RLTNQKPSFICESWYPLIKKNGMFRIMFDSVYLAFTQCFARMTLHYHQELDDY 111
DB 58 SLIDCPYPLDF-----NPLDVLWIDTXXLSA-VNCPKCMV-----LDPY 98
QY 112 FNTPGVETVPHRGSTNSLLYLNPRLKHTTGMAPLVDSLOK---LGYADETELFGARV 167
DB 99 NOTDHECKSRPDSGLSATTELDLP--GYITGPLSTYWKEMLWCVEFG-IEANAIVAVPY 155
QY 168 DFRYGLAAGHPPOVSGSK--FLKDLKNLIEASNSNNGKPVILLSHSLG---LFW--- 218
DB 156 DWRLS-----PTKEERDLVFRHKLKLTRETALKLKLG-PSIYFAISKGNVPRYILEWL 208
QY 219 -LQLNRNPPSWRKKPIKFHIALSAPWGAIDEMTYFASGNTLGVPLVPLVRDQRSS 277
DB 209 RLEIARVKLWMDQIIHAYFVAVGAPLLSGVEAKIKTSLGVTGFLP-VSEGFARLLSNSEF 267
QY 278 ESNLMLLPMPK 288
DB 268 ASLWMLMPFSK 278

RESULT 11

hypothetical protein T16H5.220 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 24-Nov-1999
C:Accession: T04767

R:Bevan, M.; De Haan, M.; Maarse, A.C.; Grievell, L.A.; Bancroft, I.; Mewes, H.W.; Mayer, submitted to the Protein Sequence Database, June 1998
A:Reference number: 215383
A:Accession: T04767

A:Molecule type: DNA

A:Residues: 1-493 <BEV>

A:Cross-references: EMBL:AL024486

A:Experimental source: cultivar Columbia; BAC clone T16H5

C:Genetics:

A:Map position: 4

A:introns: 82/1; 112/3; 152/2; 206/2; 283/3; 322/2; 340/3; 358/3

A:Note: T16H5.220

C:Superfamily: Arabidopsis thaliana hypothetical protein T16H5.220

Query Match 5.4%; Score 125.5; DB 2; Length 493;
Best Local Similarity 24.0%; Pred. No. 0.045;

Matches 50; Conservative 27; Mismatches 58; Indels 73; Gaps 10;

QY 35 NLDPILIPGNGQLEARLTNQKPTFCESWYPLIKKNG-----WFR----- 80

Db 30 NLNPVLVPGIAGSILNA-----VDHENGNERVWVRLFGADHEFR 70

QY 81 --LW--FESS-----VILAPFQCFR--ERMFLTHQELDDVFNTPGVETRPHEGSTSL 130

Db 71 TKMSRFRPSTGKTISTLDPKTSIVVPDRAGLHAIDVD-----PDM-----IVGRSV 119

QY 131 LYLNPRLKHITGYMAPLVDSLOKLGADGETLFGAPYDFRYGLAEGHPSQVSKFLKDL 190

Db 120 YFPHMIVEMIGW-----GFEGKTLFGFGYDFRQSNRLQETLDPFAKR----- 163

QY 191 KNLIEASNSNNGKPVILLSHSLGLEFV 218

Db 164 --LETVYKASGEKKINIVSHSMGGLV 188

RESULT 12

E86854 hypothetical protein oppa [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 24-Aug-2001

C:Accession: E86854

R:Bohlooti, A.; Wincker, P.; Manger, S.; Jallion, O.; Malame, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s8

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: E86854

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-600 <STO>

A:Cross-references: GB:AE005176; PID:912724867; PIDN:AAK05935.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: oppa

C:Superfamily: dipeptide transport protein

Query Match 4.8%; Score 113; DB 2; Length 600;
Best Local Similarity 21.9%; Pred. No. 0.6;

Matches 90; Conservative 58; Mismatches 159; Indels 104; Gaps 21;

QY 1 MKKQDEBLKLEVAITLVTVVVMLSILCTGASNLDPILIPGNGQLEARLTNQKP 60

Db 1 MKK-----LKV-----TLASSVLAALSLACGS-----NONSSTKRLKA 38

QY 61 STFCESWYPLIKKNGFRLMFS-----SVILAPFQCFARMTLTHQELDD-----Y 111

Db 39 GNFVAVNPNPKAIKGLKLAAYOSDSPKAEMLAPLS-----NDATGFS 83

QY 112 FNTGCVETRPVHFGSTNSLLYLPRLKHITGYMAPLVDSLOK-----LGADG 159

Db 84 MSSPG-----GGQDLFTFTNSSFKYNG--GPANSLYKDAKTATITLRKDLKMSDG 133

QY 160 ETLEGAPYDFRYGLAEGHPSQVSKFLKDLKNLIEASNSNNGKPV-----ILLSHSLG 215

Db 134 SEVTAKEYSFYDLTA--NPAYGSDRWTDLSANIV-GLSDYHAGRAKTIISITPPDGENG 190

QY 216 LEVLQILNRPNSMRKFKIKHIALSAPWGAIID--EMVTPASGVTGLGPIVLDLVDE 273

Db 191 KVIAKQFKEMTPGMNQTONGTFLEIVAPYOLKDVAPDLASAPKSTTKPLVTGPFRKN 250

QY 274 QRSSENLMLLPNPKIFGPKR-----IVIT--PIRPSAHDMDVFLDIGPPEGVY 323

Db 251 VVAGES--IKYVNPYYWG--EKPKLNSITVEIVSTAKVAALSAH-KYDYINDMRASO--- 304

QY 324 PYETRIPLIGNIKARQVPTICIMGTGVTLETIFYGKGDDDERPEISYCD 374

Db 305 -YKQ-----VKDYKGYK-----LQQLYISLIMYLGHYDKVKSISVOD 344

RESULT 13

T44608 pyruvate carboxylase (EC 6.4.1.1) [imported] - Bacillus cereus (fragment)

C:Species: Bacillus cereus

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Sep-2000

C:Accession: T44608

R:Okstad, O.A.; Hegna, I.; Lindbaek, T.; Rishovd, A.L.; Kolsto, A.B.

Microbiology 145, 621-631, 1999

A:Title: Genome organisation is not conserved between Bacillus cereus and Bacillus su

A:Reference number: 422811; MUID:99231848; PMID:10217496

A:Accession: T44608

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-984 <OKS>

A:Cross-references: EMBL:AJ010111; NID:94584147; PIDN:CAB40604.1; PID:94584148

A:Experimental source: ATCC 10987

C:Genetics:

A:Note: PYCA

C:Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl-biotin-biotin

C:keywords: ligase

Query Match 4.7%; Score 109; DB 2; Length 984;
Best Local Similarity 17.9%; Pred. No. 2.6;

Matches 104; Conservative 94; Mismatches 204; Indels 178; Gaps 24;

QY 13 VATLVTVVVMVMSILCTGASNLDPILIPGN-----GNGQLEARLTNQKPTFCES 67

Db 140 ITEMITGVIVQSOILLADGHALSKLVGPKEEVVYVHGAFQSRVTTEDPLNNFMDT 199

QY 68 WYPLIKKNGFRLMFS-----SVILAPFQCFARMT--LHYHQLDDY-----FN 113

Db 200 GKMAVYRSGGFGVRLDTGNSFGQAVITPYDLSLVKVTWALFPEQAAAKERNLKEFR 259

QY 114 TPQVETRPVHFGSTNSLLYLPRLKH--ITG-YMAPLVDS-----LQKLG 155

Db 260 IRGIKTNP-----FLENVYKHKNFLSGEVDTSPIDSPLEFLFSKRKRDGTMLN 310

QY 156 YADGETLFGAP-----YDFRYGLAEGHPSQVSKFLK----- 189

Db 311 YIGTVYVNGFPVGGKKEKPIPDARIPNVLHSEPIQKQIDDERGADGLVKWVQDKR 370

QY 190 -----LKNLIEASNSNNGKPVILLSHSLG----- 215

Db 371 VLNDTFRDAHQSLATRIKTLHQIAEPTARMLPRLFSAEMMGATPDVAYRFLKED 430

QY 216 -----LEVLQILNRPNSMRK--KFKIKHIALSAPWGAIIDMYTPASG 257

Db 431 PWERILLDERKMPNVLFQMLRSSNAVGYKNYPNLLQKFEVCSAQAG--IDVERIFDSL 488

QY 258 NTLGVPLVLDPLLVDEQSSSNLML--LNP-----KTEGP 292

Db 489 NWVEGMRAIDAVIDGTGIAETACYTGDIHDPKSKYDLVYNNLAKELEASGARILGI 548

QY 293 QKPIVITPIRPSAHDMDVFLK--IGPPEGVYPETR--ILPLIGNIKA-----POVPT 344

Db 549 KDMAGL--LKNMAVYDLVSAKETEVSIPILHATHDTSNGNGLYTTKAIENGVDLVDAVS 606

QY 345 CIMG-TGVGTLETIFYGKGDFDERPEISYGDGDTNVLVSLALQSLMKKEKNQYLKVVK 403

Db 607 SWAGTSPSANTLYALCGNERQPD-----VNIDSLERLSHYWEDRVKYAPFES 657
 QY 404 IDGVSHTSILKDEVALNEIV-GEITSINSHAE-LGSLNFP 441
 Db 658 GMAAPHT-----EYVMHEMPGGQYSNMQQAKAVGLGDRF 692

RESULT 14

B70013
 hypothetical protein yu11 - Bacillus subtilis

C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
 C:Accession: B70013

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azavedo, V.; Berter
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chc
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabre, C.; Ferrari, E.
 Nature 300, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gallier
 lech, J.; Harwood, C.R.; Hensaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
 Y. M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
 A.; Authors: Schleich, S.; Schroeter, R.; Scifone, F.; Sekiguchi, J.; Sekowska, A.; Serot
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunori, K.; Yata, K.; Yoshida, K
 A.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: B70013

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-314 <KUN>

A:Cross-references: GB:299120; GB:AL009126; MID:92635613; PIDN:CAB15191.1; PID:el184280;

A:Experimental source: strain 168

C:Genetics:

A:Gene: yu11

Query Match 4.6%; Score 106.5; DB 2; Length 314;
 Best Local Similarity 21.9%; Pred. No. 0.78;
 Matches 65; Conservative 34; Mismatches 109; Indels 89; Gaps 12;

QY 71 LKKNGWRLMFDSSVILAPFOCAEMTLHYQELDDYNTPEVRRVH----- 123
 Db 4 LVDSENHHTLRF--LKRMSYCKMEQTTPTNGTSTNAFTIPETEVMMSSRNENRT 60
 QY 124 -----FGTNSLLYLNPRLKHTGYAPVDSLOKIG 155
 Db 61 YHIFSKPSTPPAGYPIYILLDANSVGTMEAVRIGRRPEKGVIPAVYV---IG 117
 QY 156 YADGETLGCAPY-DFRYGLAAGHPQVGS-----KFL-KDLKLIEASNS 200
 Db 118 YETAEFFSARRHDEFTMPAQSKLPERPDREMPHEGAEFGFTEEDLKPEIERDYQI 177
 QY 201 NNGKXVILLSHSLGLFVQLNRP-----PS--WKKFT----KPIALSAW 244
 Db 178 DKRKQRT-FGSLGLFVQLVLLTRPDARQTYIAGSPSIHMNKPFLTKKTDHFSVLTKN 236
 QY 245 GGAID-----EMYTFASGNTLGLVPLVPLVREOR-----SSESNTMLLP 285
 Db 237 NOPINILLAGLEQHHKSRMDNARELYERLAVLSEQIRAEFCFSGEGHISVLP 293

RESULT 15

S73355

exci-nuclease ABC chain C - Mycoplasma pneumoniae (strain ATCC 29342)

N:Alternate names: hypothetical protein C09_01586; virc protein

C:Species: Mycoplasma pneumoniae

A:Variety: ATCC 29342

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Dec-1999

C:Accession: S73355

R:Himmelfreih, R.; Hilbert, H.; Plagens, H.; Pirk, E.; Li, B.C.; Herrmann, R.
 Nucleic Acids Res. 24, 4420-4449, 1996

A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumo

A:Reference number: S73327; MUID:97105885; PMID:8948633

A:Accession: S73355

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-586 <HIM>

A:Cross-references: EMBL:AE000004; GB:U00089; MID:91673671; PIDN:AAB95677.1; PID:9167

A>Note: the nucleotide sequence was submitted to the EMBL Data Library November 1996

C:Genetics:

A:Gene: uvrC

A:Genetic code: SGC3

A:Start codon: TTG

C:Superfamily: exci-nuclease ABC chain C

C:Keywords: DNA repair; hydrolase

Query Match 4.5%; Score 106; DB 1; Length 586;
 Best Local Similarity 22.2%; Pred. No. 2.1;
 Matches 106; Conservative 58; Mismatches 161; Indels 152; Gaps 24;

QY 34 SNIDPLILPGNGNOLARLTNOYRPTFCESWYPLIKKNGWRLMFDSSVILAPPT 93
 Db 64 ADVEYILKKNENDALNLEAKLIKQYKPRFNL-----VLKENNGYLYFFITASY----- 111
 QY 94 QCAEEMTL-----HYQELDDYNTPEVRRVHFGSTNSLYLNPRLK---HTGYMAP 146
 Db 112 -----KPTLELGRYEFKSKKRYFG-PFASKFRLRDIYDLKLFLPKCAPRERGHGPF 165
 QY 147 LVDLSLOKLGADSE-----TLFGAPYDFRYGLAAGHPQVGSKFLDKLNLIEAS 198
 Db 166 YVQLKMGOCMGEDTPERTQTVKGIQEFNNG-----PEQY-----LNHLQOQETKAS 215
 QY 199 NSNNGRPVILLSHSLG-GLFVQLNRPSPWKRKFTH-FIALSAPMGADIDEMTFAS 256
 Db 216 EQGNFEAA---RHFLDLQKAVLELVNMQQAFAIKAKQSHDFIG-----YVFEK 260
 QY 257 GNTLGVPL---VDPLVY-----RDEQRS-----S 277
 Db 261 -NVLAITVAAYVDNOLIGKNOQVEFLPDDEKVESALVTFIHYYSKTKIPKTLVSL 319
 QY 278 ESNLMLLPN-----PKIFGPQKPIVTPRIPYSAHMDVDFLKDIGPEGVYPERTR 329
 Db 320 EENLSTLANSKATNYQPK-NGEQKSILOTYV-----DNARYALNTKW 361
 QY 330 LPLIGNIKAPQVPTICMGVGTLETLFYGKGFDERPEISYGDGTYNIVSLALQS 389
 Db 362 TGFNNLNAEYHQALQLQVPSIQSL-----EIDISFYDKDHVG-----A 405
 QY 390 LKKEEKNOYLK-----YKIDGSHTSILKDEVALNEIVGETTSINSRAELGSLN 440
 Db 406 MLREYENGKRWKALSRRYNTINIDHKGDNTYVAD--VYVYRI---ISSIQTHKQPLPSDL 458

Search completed: February 11, 2003, 19:15:48
 Job time : 25 secs

RA Hixson J.E., Driscoll D.M., Birnbaum S., Britten M.L.:
RT "Baboon lecithin cholesterol acyltransferase (LCAT): cDNA sequences
of two alleles, evolution, and gene expression.";
RL Gene 128:295-299(1993).
CC -1- FUNCTION: CENTRAL ENZYME IN THE EXTRACELLULAR METABOLISM OF PLASMA
CC LIPOPROTEINS. AMONG OTHER SUBSTRATES IT ESTERIFIES THE FREE
CC CHOLESTEROL TRANSPORTED IN PLASMA LIPOPROTEINS.
CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + a sterol -> a sterol
CC ester + 1-acylglycerophosphocholine.
CC -1- ENZYME REGULATION: Apolipoprotein A-I IS A POTENT ACTIVATOR FOR
CC THIS ENZYME.
CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN LIVER AND CEREBELLUM.
CC -1- SIMILARITY: PARTIAL WITH LIPASES (PANCREATIC, GASTRIC, HEPATIC,
CC LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L08633; AAA35388.1; -;
DR PIR; JCI502; JCI502.
DR InterPro; IPR003386; LACT.
DR InterPro; IPR000734; Lipase.
DR Pfam; PF02450; LACT; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
KM Transferrase; Acyltransferase; Lipid metabolism; Glycoprotein; Signal.
FT CHAIN 1 24
FT SIGNAL 25 440
FT ACCT_SITE 205 205
FT DISULFID 74 98
FT DISULFID 337 380
FT CARBOHYD 44 44
FT CARBOHYD 108 108
FT CARBOHYD 296 296
FT CARBOHYD 408 408
SQ SEQUENCE 440 AA; 49645 MW; E06B0C8300BA9D9 CRC64;
Query Match 23.0%; Score 537; DB 1; Length 440;
Best Local Similarity 33.3%; Pred. No. 2.5e-35;
Matches 129; Conservative 73; Mismatches 163; Indels 22; Gaps 11;
OY PLILPGNGNOLRLNRYKPSFFICEWYPLIKKNGWRLMFDSSVILAPTCFA 97
DB 48 PYILVPGCLGNLEAKLD---KPDVY---NMW-CYRKTEDEFTTLDLNMFLPLGVDCMT 100
OY 98 ERMFLHYHQELDDYFNTPEGVETRPVPHFGSTNSLYLNPRLKHTGYMAPLVDSLOKLYA 157
DB 101 DNTRVVYNRSSGLVSNAGVQIRVGFCKTYSVEYLD--SKLAGYLTHTLVQNLNVNGV 158
OY 158 DGETLFGAPYDFRYGLAEGHPQYGSFKFLDKMLKEASNSNNGKYYILLSHSLGLF 217
DB 159 RDEYVRAAPYWRML-----EPGQ-QOEYVYHKLGLVEE-MMAAGCKPVLFGHSLGCIH 210
OY 218 VLLQNRNPPSPMRKKEIKHFTALSPAGCAIDEMVTFASNGNTLVLPDLPLVREORSS 277
DB 211 LLYFLRPOQAMKDRFIDGFTSLGAPMGSGIKPMLVLAGSDOGIPIWSSIKLKEQRIT 270
OY 278 ESNMLLNPRIFGPKQKIVITPRVSAHMDVFLDKIGPEGVYPR-ETRIILPLINNI 336
DB 271 TTSPPMFPSSRLAMPEDHYFISPSFNNTGRDQRFADLHFEQGYMMLQSR--DLLAGL 328
OY 337 KAPOVPTICMGTYGVLLETLEFYKGG-DPDERPEISYDGDGTVALVSLALQSLMKKEK 395
DB 329 PAPGEVYCLYGVGLPTPRFYIYDHGFYTDVPDLVLEDCGDTVATRS--TELCGLMGQRQ 387
OY 396 NOYLKVVKIDGVSHRSLKDEVALNEI 422
DB 388 POPVHLPLRGIQHLNMVFSNOTLEHI 414

RESULT 4
ID LCAT_MOUSE STANDARD: PRT: 438 AA.
AC P16301;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphatidylcholine-sterol acyltransferase precursor (EC 2.3.1.43)
DE (Lecithin:cholesterol acyltransferase) (Phospholipid:cholesterol
DE acyltransferase).
DE LCAT.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90094326; PubMed=2600083;
RA Warden C.H., Langner C.A., Gordon J.I., Taylor B.A., McLean J.W.,
RA Lusis A.J.;
RT "Tissue-specific expression, developmental regulation, and
RT chromosomal mapping of the lecithin: cholesterol acyltransferase
RT gene. Evidence for expression in brain and testes as well as liver.";
RL J. Biol. Chem. 264:21573-21581(1989).
RN [2]
RP SEQUENCE OF 1-14 FROM N.A.
RA Merton G., Margarett N., Taramelli R.;
RL Submitted (MAY-1992) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: CENTRAL ENZYME IN THE EXTRACELLULAR METABOLISM OF PLASMA
CC LIPOPROTEINS. AMONG OTHER SUBSTRATES IT ESTERIFIES THE FREE
CC CHOLESTEROL TRANSPORTED IN PLASMA LIPOPROTEINS.
CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + a sterol -> a sterol
CC ester + 1-acylglycerophosphocholine.
CC -1- ENZYME REGULATION: Apolipoprotein A-I IS A POTENT ACTIVATOR FOR
CC THIS ENZYME.
CC -1- SIMILARITY: PARTIAL WITH LIPASES (PANCREATIC, GASTRIC, HEPATIC,
CC LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J05154; AAA39419.1; -;
DR EMBL; X54095; CAA38029.1; -;
DR PIR; A34158; XMSN.
DR MGD; MGI:96755; Lcat.
DR InterPro; IPR003386; LACT.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF02450; LACT; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
KM Transferrase; Acyltransferase; Lipid metabolism; Glycoprotein; Signal.
FT CHAIN 1 24
FT SIGNAL 25 438
FT ACCT_SITE 205 205
FT DISULFID 74 98
FT DISULFID 337 380
FT CARBOHYD 44 44
FT CARBOHYD 108 108
FT CARBOHYD 296 296
FT CARBOHYD 397 397
FT CARBOHYD 408 408
SQ SEQUENCE 438 AA; 49765 MW; 2FD57194343136 CRC64;
Query Match 22.1%; Score 516; DB 1; Length 438;
Best Local Similarity 33.0%; Pred. No. 1.2e-33;

RL Nucleic Acids Res. 18:5308-5308(1990).
CC -1- FUNCTION: CENTRAL ENZYME IN THE EXTRACELLULAR METABOLISM OF PLASMA
CC LIPOPROTEINS. AMONG OTHER SUBSTRATES IT ESTERIFIES THE FREE
CC CHOLESTEROL TRANSPORTED IN PLASMA LIPOPROTEINS.
CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + a sterol = a sterol
CC ester + 1-acylglycerophosphocholine.
CC -1- ENZYME REGULATION: Apolipoprotein A-1 IS A POTENT ACTIVATOR FOR
CC THIS ENZYME.
CC -1- SIMILARITY: PARTIAL WITH LIPASES (PANCREATIC, GASTRIC, HEPATIC,
CC LINGULI, LIPOPROTEIN, BACTERIAL, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X54096; CAA38030.1; -.
CC PIR: S11214; XXRTN.
CC InterPro: IPR003386; LACT.
CC InterPro: IPR000734; Lipase.
CC InterPro: IPR000739; Ser_estcrs.site.
CC Pfam: PF02450; LACT; 1.
CC PROSITE: PS00120; LIPASE_SER; 1.
CC TRANSFERASE: Acyltransferase; Lipid metabolism; Glycoprotein; Signal.
CC SIGNAL 1 24
CC CHAIN 25 440 PHOSPHATIDYLCHOLINE-STEROL
CC -----
CC ACT_SITE 205 205 ACYLTRANSFERASE.
CC FT DISULFID 74 98 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT DISULFID 337 380 BY SIMILARITY.
CC FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 108 108 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 408 408 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 440 AA: 49727 MW: 65539212168A885 CRC64;

Query Match 22.0%; Score 514; DB 1; Length 440;
Best Local Similarity 32.8%; Pred. No. 1,7e-33;
Matches 128; Conservative 75; Mismatches 159; Indels 28; Gaps 12;

RESULT 7
LOCAT_TATKG STANDARD: PRT; 293 AA.
AC 035840.
ID LOCAT_TATKG
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphatidylcholine-sterol acyltransferase (EC 2.3.1.43) (Lecithin-
DE cholesterol acyltransferase) (Phospholipid-cholesterol
DE acyltransferase) (Fragment).
GN LCAT.
OS Tatera kempii gambiana (Kemp's gerbil).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
OC Tatera.
OX NCBI_TaxID=41264;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98086782; PubMed=9417899;
RA Robinson M., Catzeffis F., Briolay J., Mouchiroud D.;
RT "Molecular phylogeny of rodents, with special emphasis on murids:
RT evidence from nuclear gene LCAT".
RL Mol. Phylogenet. Evol. 8:423-434(1997).
CC -1- FUNCTION: CENTRAL ENZYME IN THE EXTRACELLULAR METABOLISM OF PLASMA
CC LIPOPROTEINS. AMONG OTHER SUBSTRATES IT ESTERIFIES THE FREE
CC CHOLESTEROL TRANSPORTED IN PLASMA LIPOPROTEINS.
CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + a sterol = a sterol
CC ester + 1-acylglycerophosphocholine.
CC -1- ENZYME REGULATION: Apolipoprotein A-1 IS A POTENT ACTIVATOR FOR
CC THIS ENZYME (BY SIMILARITY).
CC -1- SIMILARITY: PARTIAL WITH LIPASES (PANCREATIC, GASTRIC, HEPATIC,
CC LINGULI, LIPOPROTEIN, BACTERIAL, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U72298; AAB58989.1; -.
CC DR EMBL: U72297; AAB58989.1; JOINED.
CC DR InterPro: IPR003386; LACT.
CC DR InterPro: IPR000734; Lipase.
CC DR Pfam: PF02450; LACT; 1.
CC DR PROSITE: PS00120; LIPASE_SER; 1.
CC TRANSFERASE: Acyltransferase; Lipid metabolism; Glycoprotein.
CC NON_TER 1 1
CC ACT_SITE 123 123 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT NON_TER 293 293
CC SEQUENCE 293 AA: 33822 MW: 86CFB4BB7C6A738B7 CRC64;

Query Match 15.2%; Score 354; DB 1; Length 293;
Best Local Similarity 29.2%; Pred. No. 5.6e-21;
Matches 98; Conservative 57; Mismatches 129; Indels 52; Gaps 10;

```

Db 137 GIPIMSSIKREBORITTTSPMWFPRDVPEDHVFISTPEFNVTGODEFRFSDLHFE 196
Qy 321 GYPPY-ETRIPLPLIGNIKAPQVITTCMGVGTLETLFYKGDPDEREIS--YDDGSG 377
Db 197 GYMWYFLOS--DLLEGLPAPGVVCLYGVGRPTRYIYDHNFPYKDPVALYEDGD 253
Qy 378 TVNLVSLALQSLMKKEKNQYLKVKIDGVSHTSL 413
Db 254 TVATRS--TELCGOMGROSQPVYLLPMNGTDHLMV 288

RESULT 8
LCAT_MICM  STANDARD:  PRT:  299 AA.
ID LCAT_MICM
AC 035724:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphatidylcholine-sterol acyltransferase (EC 2.3.1.43) (Lecithin-
DE cholesterol acyltransferase) (Phospholipid-cholesterol
DE acyltransferase) (Fragment).
GN LCAT.
OS Micromys minutus (European harvest mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC Micromys.
OX NCBI_Taxid=13151;
RN SEQUENCE FROM N.A.
RA MEDLINE-98086782; PubMed-9417899;
RX Robinson M., Catzeffis F., Briolay J., Mouchiroud D.;
RT "Molecular phylogeny of rodents, with special emphasis on murids:
RT evidence from nuclear gene LCAT."
RL Mol. Phylogenet. Evol. 8:423-434(1997).
CC -1- FUNCTION: CENTRAL ENZYME IN THE EXTRACELLULAR METABOLISM OF PLASMA
CC LIPOPROTEINS. AMONG OTHER SUBSTRATES IT ESTERIFIES THE FREE
CC CHOLESTEROL TRANSPORTED IN PLASMA LIPOPROTEINS.
CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + a sterol = a sterol
CC ester + 1-acylglycerophosphocholine.
CC -1- ENZYME REGULATION: APOLIPOPROTEIN A-I IS A POTENT ACTIVATOR FOR
CC THIS ENZYME (BY SIMILARITY).
CC -1- SIMILARITY: PARTIAL WITH LIPASES (PANCREATIC, GASTRIC, HEPATIC,
CC LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sib.ch).
CC -----
CC EMBL: U72294; AAB58988.1; -.
CC DR EMBL: U72293; AAB58988.1; JOINED.
CC DR InterPro: IPR003386; LACT.
CC DR InterPro: IPR000734; Lipase.
CC DR Pfam: PF02450; LACT; 1.
CC DR PROSITE: PS00120; LIPASE_SER; FALSE_NEG.
CC TRANSFERASE; Acyltransferase; Lipid metabolism; Glycoprotein.
CC KM
CC FT NON_TER 1
CC FT CAROHD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CAROHD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CAROHD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CAROHD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT NON_TER 299
CC SQ SEQUENCE 299 AA; 34833 MW; 55896C9A38C2482C CRC64;

Query Match 14.7%; Score 342.5; DB 1; Length 299;
Best Local Similarity 27.8%; Pred. No. 4.7e-20;
Matches 94; Conservative 61; Mismatches 136; Indels 47; Gaps 10;
Qy 78 WFLMFSSVILAPFTQCFAERWTLHYHQELDDYFMTPGVETRVPHFGSTNSLLYNPL 137

```

```

Db 1 FETIMIDINMFLPLGVDCWIDNTNRVYNNSSGRMSNAPQVQIWPFGKTYSEYILD-- 58
Qy 138 KHTGTMAPLVDSLOKLGADGETLFGARYDPRTYGLAAGHPQVSKELKDLKNIIEA 197
Db 59 NKIAGYLHLYLVONLVNNAAYRDRVAPDWRLA-----PSQ-ODEYYOKIAELVEEM 111
Qy 198 SNSNCKRPYLLSHSGGLEVLQDLNRNPPSKRFKIFHIALSADMGAIIDEMYFASG 257
Db 112 YDA-YKRPVFLIHRIGCLHVLHFLHQ--SMK----- 141
Qy 258 NFLGVLVDPLVLRDORSSESNMLLPPNPKIFGPQKPIYIPRYSADHWDFLKDG 317
Db 142 --GIPIMSNIKREBORITTTSPMWFPAHVVEDHVFISTPEFNVTGODFFKRPADLH 198
Qy 318 FPEGYPPY-ETRIPLPLIGNIKAPQVITTCMGVGTLETLFYKGK-DEDEREISYGG 375
Db 199 FEKGWYFLOS--DLLEGLPAPGVVCLYGVGRPTRYIYDHNFPYKDPVALYEDG 256
Qy 376 DGTVNLVSLALQSLMKKEKNQYLKVKIDGVSHTSL 413
Db 257 DETVATRS--TELCGOMGROSQPVYLLPMNGTDHLMV 293

RESULT 9
LCAT_ELIOU  STANDARD:  PRT:  299 AA.
ID LCAT_ELIOU
AC 03573:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphatidylcholine-sterol acyltransferase (EC 2.3.1.43) (Lecithin-
DE cholesterol acyltransferase) (Phospholipid-cholesterol
DE acyltransferase) (Fragment).
GN LCAT.
OS Elionys merctus (Garden dormouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Myoxidae; Leithinae;
OC Elionys.
OX NCBI_Taxid=53277;
RN SEQUENCE FROM N.A.
RA MEDLINE-98086782; PubMed-9417899;
RX Robinson M., Catzeffis F., Briolay J., Mouchiroud D.;
RT "Molecular phylogeny of rodents, with special emphasis on murids:
RT evidence from nuclear gene LCAT."
RL Mol. Phylogenet. Evol. 8:423-434(1997).
CC -1- FUNCTION: CENTRAL ENZYME IN THE EXTRACELLULAR METABOLISM OF PLASMA
CC LIPOPROTEINS. AMONG OTHER SUBSTRATES IT ESTERIFIES THE FREE
CC CHOLESTEROL TRANSPORTED IN PLASMA LIPOPROTEINS.
CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + a sterol = a sterol
CC ester + 1-acylglycerophosphocholine.
CC -1- ENZYME REGULATION: APOLIPOPROTEIN A-I IS A POTENT ACTIVATOR FOR
CC THIS ENZYME (BY SIMILARITY).
CC -1- SIMILARITY: PARTIAL WITH LIPASES (PANCREATIC, GASTRIC, HEPATIC,
CC LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sib.ch).
CC -----
CC EMBL: U72316; AAB58999.1; -.
CC DR EMBL: U72315; AAB58999.1; JOINED.
CC DR InterPro: IPR003386; LACT.
CC DR InterPro: IPR000734; Lipase.
CC DR Pfam: PF02450; LACT; 1.
CC DR PROSITE: PS00120; LIPASE_SER; FALSE_NEG.
CC TRANSFERASE; Acyltransferase; Lipid metabolism; Glycoprotein.
CC KM
CC FT NON_TER 1

```

FT ACT_SITE 127 127 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 299 299
 SO SEQUENCE 299 AA; 34653 MW; 9C2041F7E79A2CEP CRC64;
 Query Match 14.6%; Score 341; DB 1; Length 299;
 Best Local Similarity 27.6%; Pred. No. 6,2e-20;
 Matches 93; Conservative 59; Mismatches 137; Indels 48; Gaps 9;
 QY 79 FRLMFDSSVILAPFTQCFAERMTLHYHQELDDYFNTPGVETRPVPHGSGTSLNLYLNPRLK 138
 DB 4 FTIMLDMILFSLGVDICIDINTRVYVNRSSGYSNAPQOIRVPGFKTYSVEYLD--N 61
 QY 139 HITGMAPLVDSLOKTLGADGETLFGAPYDFRYGLAEGHPQVSGSKFLDKNLIEEAS 198
 DB 62 KLAGMHTLVQNLVNNKAVRDETVAPRYDMRL-----EPHQOEYILK-LAGLVEE-M 113
 QY 199 NSNNGRPVILLSHSLGGLFVQLNLNRPSPWRKKFIKHFIALSAPWGALIDEMYTASGN 258
 DB 114 YATYGRPFLLGHSLGFCCHLLYFLLLQPO----- 142
 QY 259 TLGVLVPLVLRDQBSSESNTMLLPNPKITFGPKPIYITPIRYSAHMDVFLKDIGF 318
 DB 143 --GIPIKSSIKLVEQRIITTSPPMFPHQWPEDEHVFISTPNNFYTESDFQRFADLAF 200
 QY 319 PEGVYPR-ETRIPLPLGNIRKAPQVPTICMGTVGTLETLYFGKG-DDEPEREISGDG 376
 DB 201 EDGWMMLQGR--DLGLPLPAPGVEYCLYGVLPTPTHTYMDHGFPTTDPVGIITDEGD 258
 QY 377 GTVNLVSLALQSLMKKEKNQYLKVKIKIDGVSHSIL 413
 DB 259 DTVTTHS-IELCSHWQGRQPOPHVLLPLRGTOHLNMY 294
 RESULT 10
 LCAT_CUEGL STANDARD: PRT: 291 AA.
 AC 035502;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylcholine-sterol acyltransferase (EC 2.3.1.43) (Lecithin-
 cholesterol acyltransferase) (Phospholipid-cholesterol
 acyltransferase) (Fragment).
 GN LCAT.
 OS Caethlonomys glareolus (Bank vole).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Arvicolinae;
 OC Clethrionomys.
 OX NCBI_TaxID=51090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98086782; PubMed=9417899;
 RA Robinson M., Cacaceffis F., Briolay J., Mouchiroud D.;
 RT "Molecular phylogeny of rodents, with special emphasis on murids:
 RT evidence from nuclear gene LCAT".
 RL Mol. Phylogenet. Evol. 8:423-434(1997).
 CC -1- FUNCTION: CENTRAL ENZYME IN THE EXTRACELLULAR METABOLISM OF PLASMA
 CC LIPOPROTEINS. AMONG OTHER SUBSTRATES IT ESTERIFIES THE FREE
 CC CHOLESTEROL TRANSPORTED IN PLASMA LIPOPROTEINS.
 CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + a sterol - a sterol
 CC ester + 1-acylglycerophosphocholine.
 CC -1- ENZYME REGULATION: APOLIPOPROTEIN A-I IS A POTENT ACTIVATOR FOR
 CC THIS ENZYME (BY SIMILARITY).
 CC -1- SIMILARITY: PARTIAL WITH LIPASES (PANCREATIC, GASTRIC, HEPATIC,
 CC LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U72300; AAB58990.1; JOINED.
 DR EMBL; U72299; AAB58990.1; JOINED.
 DR InterPro; IPR003386; LACT.
 DR InterPro; IPR000734; Lipase.
 DR Pfam; PF02450; LACT; 1.
 DR PROSITE; PS00120; LIPASE_SER; 1.
 KW Transferase; Acyltransferase; Lipid metabolism; Glycoprotein.
 FT NON_TER 1 1
 FT ACT_SITE 125 125 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 291 291
 SO SEQUENCE 291 AA; 33536 MW; B1F7E9DB8201219D CRC64;
 Query Match 13.9%; Score 323.5; DB 1; Length 291;
 Best Local Similarity 28.5%; Pred. No. 1.5e-18;
 Matches 87; Conservative 53; Mismatches 112; Indels 53; Gaps 9;
 QY 78 WFLMFDSSVILAPFTQCFAERMTLHYHQELDDYFNTPGVETRPVPHGSGTSLNLYLNPRL 137
 DB 1 FFTFLDMFLMVLGVCNMDINTRVYVNRSSGRMSNAPQOIRVPGFKTYSVEYLD-- 58
 QY 138 HITGMAPLVDSLOKTLGADGETLFGAPYDFRYGLAEGHPQVSGSKFLDKNLIEEA 197
 DB 59 NKLACGMHTLVQNLVNNKAVRDETVAPRYDMRL-----EPSQ-QEYIKLAGLVEE 110
 QY 198 NSNNGRPVILLSHSLGGLFVQLNLNRPSPWRKKFIKHFIALSAPWGALIDEMYTASG 257
 DB 111 MHAAGKPFVFLGHSGVGLHLVLF----- 134
 QY 258 NTLGVLVPLVLRDQBSSESNTMLLPNPKITFGPKPIYITPIRYSAHMDVFLKDIG 317
 DB 135 -NGIPIKSSIKLREQRIITTSPPMFPARVPEDEHVFISTPNNFYTESDFQRFADLAF 193
 QY 318 PEGVYPR-ETRIPLPLGNIRKAPQVPTICMGTVGTLETLYFGKGDEPEREIS--YGD 374
 DB 194 FEEGWMMLQGR--DLGLPLPAPGVEYCLYGVLPTPTHTYMDHGFPTTDPVGIITDEGD 250
 QY 375 GDGYV 379
 DB 251 GDDPVT 255
 RESULT 11
 PDAT_SCHPO STANDARD: PRT: 623 AA.
 ID PDAT_SCHPO
 AC 094680;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Potential phospholipid:diacylglycerol acyltransferase (EC 2.3.1.158)
 DE (PDAT).
 GN SPBC776.14.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21846401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., Mclean J.,

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Obell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbhnowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymompres B.,
 RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Nuer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
 RA Corfau A., Cadieu E., Diano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovskii G.V., Usseery D., Barrell B.G., Nurse P.,
 RT The genome sequence of *Schizosaccharomyces pombe*.
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: Triacylglycerol formation by an acyl-CoA independent
 CC pathway. The enzyme specifically transfers acyl groups from the
 CC sn-2 position of a phospholipid to diacylglycerol, thus forming
 CC an sn-1-lysophospholipid (by similarity).
 CC -1- CATALYTIC ACTIVITY: Phospholipid + 1,2-diacylglycerol =
 CC lysophospholipid + triacylglycerol.
 CC -1- SIMILARITY: SOME, TO MAMMALIAN PHOSPHATIDYLCHOLINE-STEROL O-
 CC ACYLTRANSFERASE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL035263; CAA22887.1; -
 CC DR InterPro: IPR003386; LACT.
 CC Pfam: PF02450; LACT; 1.
 CC KM Hypothetical protein; Transferase; Acyltransferase; Transmembrane.
 CC FT DOMAIN 11 14 POLY-LYS.
 CC TRANSMEM 57 77 POTENTIAL.
 CC SEQ SEQUENCE 623 AA; 69748 MW; 1140ED8D46968C90 CRC64;
 CC -----
 CC Query Match 7.8%; Score 181; DB 1; Length 623;
 CC Best Local Similarity 22.1%; Pred. No. 9.7e-07;
 CC Matches 109; Conservative 65; Mismatches 171; Indels 148; Gaps 24;
 CC -----
 CC 38 PLILIPNGNGNLEARLITNOYKPSFTICESW-----YPLIKRKNQWFRILMPDSSVILAP 91
 CC 147 PVLIVPGVIVSSGL-----ESMSFNNGCSIPYFRK-----RLMGSMMLKAM 166
 CC 92 F--TQCFAEAMTLHYHOELDDYNTPEVETRVPH-FGSTNSLLYLNRKLHITGY--MAP 146
 CC 187 FLDRKQCMLEHMLMDKRTGLD---PKGIKRLAAQGEAAAF-----ITGYTWMSK 233
 CC 147 LVDSLKLGYADGETLEGAYDFRYGLAAGHPSOYGSKTLKDKLNLEASNNNGKPY 206
 CC 234 VINELAIGT-EPNNMLSASTYDKRLSYA---NLEERKQTFPSKLMFT-EYSNVYHKKV 287
 CC 207 ILLSHSLGGLFVLQL-----NRNPSMRKRKFKHFKIALSAPMGAIIDEMTFEASG- 257
 CC 288 VLLSHSGSGOVYTYFFPKWAEAGYNGGPGTWVDHIEAFINIGSSGLGAKRTVAALLSGE 347
 CC 258 -NVLGVPVLPVLLVRDEQSSSESNL-----WLPNPKIFGPO----- 293
 CC 348 MKDGTGIVITLNLKEFFSRERAMVMTGWSMPLPKGADVAPDDLQNTFNSGALIRY 407
 CC 294 KPIVIRPIREYSADWDVDFKDI-----GPPEGVYPTETRL-----P 331
 CC 408 REPIDDKHDEFDDIALQFLKNTVDDDFKYLAKNYSHGILAMTEKEVELKNNEMPSKWINP 467
 CC 332 LIGNIK-APQVPITCIMGTVGTLFTLLFYGKGFDERPE-----ISY 372

DB 468 LERSLPAAPDMKITYCHGVGKPTERGYXY-----TNNPEGQPIDSSVNDGTRKENGIVW 522
 CC 373 GDDGDEVNLYSL-LALQSLMKKEKNQYKLVKIDGYSHTSI-----LKDEVALNEIYG---- 424
 CC 523 DDDGDTPLPLIALGLVCKRWQ-----TKRNPANPNTSITWYKHEPAADLGGPR 573
 CC 425 -EITSINSHAEI 435
 CC 574 SAEHVILHSEL 586
 CC -----
 CC RESULT 12
 CC UVRC_MYCPN STANDARD; PRT; 586 AA.
 CC ID UVRC_MYCPN
 CC AC P75350;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Excluclease ABC subunit C.
 CC GN UVRC OR MPN125 OR MP029.
 CC OS Mycoplasma pneumoniae.
 CC OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 CC OX NCBI_TaxID=2104;
 CC [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-ATCC 29342 / M129;
 CC RX MEDLINE=97105885; PubMed=8948633;
 CC RA Haimelreich R., Halbert H., Plagens H., Pirkl E., Li B.-C.,
 CC RA Hermann R.;
 CC RT "Complete sequence analysis of the genome of the bacterium *Mycoplasma*
 CC pneumoniae";
 CC Nucleic Acids Res. 24:4420-4449(1996)
 CC CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
 CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
 CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). ATTACHES TO THE
 CC UVRA-UVRB COMPLEX, DISPLACING UVRA, AND THE DAMAGED DNA STRAND IS
 CC NICKED ON BOTH SIDES OF THE DAMAGED SITE (BY SIMILARITY).
 CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS; UVRA, UVRB AND UVRC.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE UVRC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE000004; AAB95677.1; -
 CC DR InterPro: IPR004791; UVRC.
 CC DR InterPro: IPR001162; UVRC_C.
 CC DR InterPro: IPR000305; UVRC_N.
 CC DR Pfam: PF01541; Excl_endo_N; 1.
 CC DR ProDom: PD005870; UVRC_C; 1.
 CC DR SMART: SM00465; GYIC; 1.
 CC DR TIGRfam: TIGR00194; uvrc; 1.
 CC KW SOS response; Excision nuclease; DNA repair; Complete proteome.
 CC SEQ SEQUENCE 586 AA; 67771 MW; 6A495512C3447C81 CRC64;
 CC -----
 CC Query Match 4.5%; Score 106; DB 1; Length 586;
 CC Best Local Similarity 22.2%; Pred. No. 0.86;
 CC Matches 106; Conservative 58; Mismatches 161; Indels 152; Gaps 24;
 CC -----
 CC 34 SNLDPLILIPNGNGNLEARLITNOYKPSFTICESWYPLIKRKNQWFRILMPDSSVILAPFT 93
 CC 64 ADVEYVILKNDALNLEAKLIKQYKRFNL-----VLKENNGYLYFFITASY----- 111
 CC 94 QCFAEKRTL-----HYHOELDDYNTPEVETRVPHFGSTNSLLYLNRKL---HITGYMAP 146
 CC 112 -----KPTLGLGRYERSKKNRYFG-PPASSKFRILRDYILLKLFLPKCAKAPHERHPCP 165


```

OS Rhizobium sp. (strain NGR234).
OC Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freilberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RT Perret X.;
RL "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- SIMILARITY: NONE OBVIOUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AE000081; AB91737.1;
DR InterPro; IPR000379; Set-estrs-site.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 434 AA; 47425 MW; BA0FCCE3B2E494B CRC64;

Query Match
Best Local Similarity 4.3%; Score 99.5; DB 1; Length 434;
Matches 50; Conservative 41; Mismatches 77; Indels 65; Gaps 12;

OY 115 PGEVRVPHFGSTNSLLYINPRLKHTGYMAPLVDSLOKIGYADGRTLGAPYDFRYGLA 174
DB 173 PADKKLETFG-----LSGWLASTKEHALALGSAAPADATTA-----GLL 211
OY 175 ASGHPS--QVSGSKFLKD---LKN-----LIE-----EASNSNNGKPYLLS 210
DB 212 ALGRRSAHQKASRFLGDVYVLYLNKRTAAAPGIYRAVLQAFRNDAARQEGDKDLITVG 271
OY 211 HSLGGLFVQLIDNRPSPRRKFKIKHILASAPWGALIDEMTFASGNTLGVPLVPLV 270
DB 272 HSLGGVITYDILSYRAPDLK---VDYLATIGSQVGLFDEMLIYSRPLPPGMP-TDP--- 323
OY 271 RDEORSSSENLT--WLLPNPKIFGPOKPIYITPIRPSAHDMVDFLKDIFGPEG 321
DB 324 KMDRLAANPNVGKWL---NVFD-----TNDYFESSAGIYKSVADPQYDTG 366

RESULT 15
ETFL_FOMPV STANDARD; PRT; 633 AA.
AC P21966; O9J5F2;
DT 01-AUG-1991 (Rel. 19, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Early transcription factor 70 kDa subunit (VETf small subunit).
GN FPOV057 OR FPOD6.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=FP-1;
RX MEDLINE=90324937; PubMed=2165135;
RA Tataraglia J., Winslow J., Goebel S., Johnson G.P., Taylor J.,
RA Paoletti E.;
RL "Nucleotide sequence analysis of a 10.5 kbp HindIII fragment of
RT fowlpox virus: relatedness to the central portion of the vaccinia
RL virus HindIII D region.";
RL J. Gen. Virol. 71:1517-1524(1990).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=20193820; PubMed=10729156;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RL "The genome of fowlpox virus.";
RL J. Virol. 74:3815-3831(2000).
CC -1- FUNCTION: ACTS WITH RNA POLYMERASE TO INITIATE TRANSCRIPTION FROM
CC EARLY GENE PROMOTERS. A DNA-DEPENDENT APPASE ACTIVITY IS
CC ASSOCIATED WITH VETf.
CC -1- SUBUNIT: HETERODIMER OF A 70 KDA AND A 82 KDA SUBUNIT.
CC -1- SIMILARITY: BELONGS TO THE VETf SUBFAMILY OF HELICASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X17202; CA35069.1;
DR EMBL; AF198100; AAF4401.1;
DR PIR; F35216; F35216.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002464; DEAD_Box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00690; DEAD_ATP_HELICASE; 1.
KW Helicase; Transcription regulation; Activator; ATP-binding.
FT NP_BIND 45
FT SITE 135 138 DEXH BOX.
FT FT CONFLICT 605 633
FT REF. 1).
SQ SEQUENCE 633 AA; 73020 MW; C17DE6A97EBEAD25 CRC64;

```

```

Query Match
Best Local Similarity 4.2%; Score 99; DB 1; Length 633;
Matches 89; Conservative 58; Mismatches 144; Indels 138; Gaps 23;

OY 71 LIRKKNQWFLMFDSSVILIAPFQCAEEMTLHYHELDYFNTPEVERVPHFGSTNSL 130
DB 215 ILNKKISTYEL-----COTELPSLTFHK--EFLDTKYVYCKMSKLOETD-- 257
OY 131 LYLNPRL-----KHIT-----GYMAP--LYVDSLOKIGYADGETLF-----GAPYDF 169
DB 258 -YINVRKLCNNEMFEKMNINVSILAVGPLNANSLELLEVEQDKELYPLMKINDVLY-- 314
OY 170 RIGLAEGHPQVSGSKFLDKLNLIEASNSNNGKPYILLSHS-LGGLFVQLIDNRPN--- 225
DB 315 ---GDELTKLNISKF---KYFIDTIGNL--TGKNFIYFSNSTYGGIVIKYIMLNNGYS 365
OY 226 -----PPSWKRRKFKIKHILASAPWGALID---EMYTFASGNTLGVPLVPLVRDQ 274
DB 366 EYASOGQINPKMINKHLKTFALVTSKMSKSLLELVYN-SPENDNGEKI---WFLPSSN 421
OY 275 RSSES-----NLMLLPNPKIFGPOKPIYITPIRPSAHDMVDFLKDIFGPEGYEPT 327
DB 422 IMSESTLIKVENIMWMITPDRFSQYNQILGRIRKFS-----YKDVSKPYNRY--- 470
OY 328 RILPLIGNIKAP-QVPIYICMGTVGTLTFLYFGKDPDERPEISYGGDGGVYVNLVSLA 386
DB 471 ---LLATVYSPENDTITSIDYSIDINTL---PFD-----IKLLLY 506
OY 387 LOSLWKEEKNQVLYKVIDGVSHTSLTKD-----EVALNEIFYGITSINSH 432
DB 507 LKFKTKETKRIY-----STLKDLSTNYRSSPHPIYITDVVLGELVROFPYHNRS 554
OY 433 AELGLSNLF 441
DB 555 VSINDAKLF 563

Search completed: February 11, 2003, 19:14:37
Job time : 17 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 19:12:56 : Search time 35 Seconds
(without alignments)
2607.968 Million cell updates/sec

Title: US-09-857-612A-14
2333
Perfect score: 1 MKRQEGSLKIEVATLVTV.....GEITSINSHAEIGLSNLEFSG 443
Sequence:

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1380.5	59.2	432	10 Q9FZ18	Q9FZ18 arabidopsis
2	831	35.6	822	10 Q8S5X5	Q8S5X5 oryza sativ
3	792.5	34.0	266	10 Q9SDR9	Q9SDR9 prunus dulc
4	673.5	28.9	363	10 Q8RYV9	Q8RYV9 oryza sativ
5	612	26.2	412	11 Q8VER4	Q8VER4 mus musculu
6	610	26.1	412	4 Q9Y2B3	Q9Y2B3 homo sapien
7	586	25.1	407	6 Q8WMP9	Q8WMP9 bos taurus
8	570	24.4	421	5 Q9Y168	Q9Y168 drosophila
9	560	24.0	421	5 Q9Y179	Q9Y179 drosophila
10	546	23.4	440	6 Q9MZ04	Q9MZ04 tupia glis
11	533.5	22.9	451	13 Q9DDJ6	Q9DDJ6 anas platyr
12	526	22.5	440	11 Q35849	Q35849 ratius norv
13	518	22.0	444	11 Q9CWA7	Q9CWA7 mus musculu
14	512.5	22.0	417	5 Q21515	Q21515 caenorhabdi
15	457.5	19.6	272	4 Q9UG04	Q9UG04 homo sapien
16	436.5	18.7	443	10 Q94HP5	Q94HP5 oryza sativ

17	369.5	15.8	298	11 Q35726	Q35726 marmota mar
18	364	15.6	307	11 Q35503	Q35503 cricetus
19	351.5	15.1	291	11 Q35746	Q35746 spalax leuc
20	351	15.0	295	11 Q35574	Q35574 gerbillus h
21	342	14.7	294	11 Q54727	Q54727 akodon torq
22	340	14.6	297	11 Q35591	Q35591 myoxus glis
23	336	14.4	296	11 Q35747	Q35747 nanospalax
24	331	14.2	292	11 Q35839	Q35839 sciurus gril
25	329.5	14.1	274	11 Q35590	Q35590 myocastor c
26	325.5	14.0	268	11 Q9EP2	Q9EP2 dipus sagit
27	323.5	13.9	268	11 Q9EP6	Q9EP6 allactaga e
28	322.5	13.8	268	11 Q9EP5	Q9EP5 calomyscus
29	319.5	13.7	268	11 Q9EP0	Q9EP0 phodopus ro
30	317.5	13.6	272	11 Q35830	Q35830 rhizomys pr
31	315.5	13.5	268	11 Q9EP9	Q9EP9 lophuromys
32	314.5	13.5	267	11 Q9EP0	Q9EP0 jaculus jac
33	311.5	13.4	268	11 Q9EP8	Q9EP8 mystromys a
34	307	13.2	256	11 Q9EPC9	Q9EPC9 sicista kaz
35	305.5	13.1	268	11 Q9EP3	Q9EP3 dendromys m
36	304.5	13.1	268	11 Q9EP4	Q9EP4 neotoma fus
37	301.5	12.9	268	11 Q9EP3	Q9EP3 napaeozapus
38	300.5	12.9	192	4 Q9NP06	Q9NP06 homo sapien
39	299.5	12.8	268	11 Q9EP1	Q9EP1 dicystomys
40	299.5	12.8	268	11 Q91YH4	Q91YH4 cricetomys
41	298.5	12.8	265	11 Q91YB3	Q91YB3 sacostomus
42	295.5	12.7	264	11 Q9EP6	Q9EP6 macrostom
43	295.5	12.7	268	11 Q9EP5	Q9EP5 myospalax s
44	294.5	12.6	268	11 Q9EP2	Q9EP2 nesomys rut
45	285.5	12.2	266	11 Q9EPC8	Q9EPC8 steatomys s

ALIGNMENTS

RESULT 1
Q9FZ18 PRELIMINARY; PRT; 432 AA.
AC Q9FZ18;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE F17L21.27 (F17L21.27 (F17L21.27 (F17L21.27, Last annotation update)
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Johnson-Hopson C., Brooks S., Buehler E., Chao Q., Khan S., Kim C.,
RA Shinn P., Altafi H., Bel Q., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharzky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F17L21 from chromosome
I.";
RT Submitted (Apr-1998) to the EMBL/Genbank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RT Submitted (Apr-1998) to the EMBL/Genbank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RT Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RT Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
[5]
RP SEQUENCE FROM N.A.

RA Ecker J.R.:
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altai H., Bei B., Chin C., Chlou J., Choi E.,
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharly N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thavert A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.,
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,
 RA Bosser L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jones T., Kamliya A., Kallin-Neumann G., Kawai J., Lam B.,
 RA Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,
 RA Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,
 RA Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K.,
 RA Davis R.W., Theologis A., Ecker J.R.:
 RT "Arabidopsis cDNA clones."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC004557; AA99739.1; -
 DR EMBL; AF367336; AA32913.1; -
 DR InterPro; IPR003386; LACT.
 DR InterPro; IPR003379; Ser_estrs_site.
 DR Pfam; PF02450; LACT.1.
 SO SEQUENCE 432 AA; 48307 MW; 4ACD9608C6785078 CRC64;

Query Match 59.2%; Score 1380.5; DB 10; Length 432;
 Best Local Similarity 61.0%; Pred. No. 1.7e-111;
 Matches 264; Conservative 62; Mismatches 92; Indels 15; Gaps 5;

QY 1 MKKDEESLKEVNTLVTVVVMVSLCTCGASVLDPLILIPGNGQLEARLTNOYKP 60
 DB 1 MKK-----TSSHYSVAILLVVVTMTSMCAVGSNVPLVLPVGNGQLEARDREYKP 55
 QY 61 STFCESW-YPLIKKNGMFRMFDSVILAFPTQCFARMTLHYHQELDVFNTPGVET 119
 DB 56 SSVCMSSWLYPIHKSGCMFRMFDAVALLSPFTRCFSDRMALVYDPLDDYQVNAFGVOT 115
 QY 120 RVPHFGSTNSLLYNPRKLTGTAKPVLVDSLQ-KLGYADEGTLFGADYDRYGLAAGH 178
 DB 116 RVPHFGSTKSLLYLDRPRLDATSMEHLVKALEKCGVNDOTILGADYDRYGLAAGH 175
 QY 179 PSQVSKFLDKLNLIEASNSNNGKPYILLSHSLGSLFVQLNLRNPPSRKKFKIKFI 238
 DB 176 PSRAVSQFLQDLQKLVETSSSENEKRPYILLSHSLGSLFVQLNLRNPPSRKKFKIKFI 235
 QY 239 ALSAPMGGAIDEMVTFASGNTLVGPVLDPLVLRDQSSSESNLMLPNPKIFGPO-KPIV 297
 DB 236 ALAAPMGGTISQMTKTFASGNTLVGPVLDPLVLRHQRTSESNOMLPTKVFHDRTKPLV 295
 QY 298 ITPRPYSAHMVDPLKIGFPEGVYPERILPLIGNIKAPQVPTITMGVGTLETL 357
 DB 296 VTPQVNTAYVEMDRFDLIGFSOGVVPYKTRVPLPLETELMPGVAVTCIYGRGVDTPEVL 355
 QY 358 FYGKGDPERPEISYGGDGTVNLYSLALQSLMKEEKQYLLKVKYIDGSHTSILKDEY 417
 DB 356 MYGKGGEFKQPEIKYGGDGTVNLYSLALQSLMKEEKQYLLKVKYIDGSHTSILKDEY 408
 QY 418 ALNEIVGEITSIN 430
 DB 409 ALKEIMKQISTIN 421

RESULT 2
 Q853X5 PRELIMINARY; PRT; 822 AA.
 AC Q853X5;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Hypothetical 91.2 kDa protein.
 GN O0175C11.13.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Euphorbiaceae; Oryzaeae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
 RA Saeki C., Henry D., Oates R., Simmons J.;
 RT "Rice Genomic Sequence."
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC103891; AAM19133.1; -
 KW Hypothetical protein.
 SO SEQUENCE 822 AA; 91220 MW; 8F329D21E6232ABE CRC64;

Query Match 35.6%; Score 831; DB 10; Length 822;
 Best Local Similarity 42.2%; Pred. No. 2.8e-63;
 Matches 183; Conservative 70; Mismatches 151; Indels 30; Gaps 10;

QY 18 VTVVVVHLSL-----LCTCGASNDPLILIPGNGQLEARLTNOYKSTFCESWYPLIK 73
 DB 395 VAVVAVVLSLPSRSGSGGSDLHPVYLVPGYGNRLYARLTAAVEPAAPRCGA----RE 450
 QY 74 KKNQMFRLM-FDSSVILAFPTQCFARMTLHYHQELDVFNTPGVETRVPHFGSTNSLY 132
 DB 451 GKDEWFQMPIDAAASEPAQAPCLAEMKSLYVDVADYDRYVAGVYTRVPSFASALV- 509
 QY 133 LNPRIKHTGYMADLVSLQKLGYADEGTLFGADYDRYGLAAGHPSQVSKFLDKLN 192
 DB 510 -----GW-DPLVQLQLEAMGHRDGSGLFAADYDRYVAVAPRHPSAAGERYFARLTR 559
 QY 193 LIEBASNSNNGKPYILLSHSLGSLFVQLNLRNPPSRKKFKIKFIKLSAPMGGAIDEMV 252
 DB 560 LIEBASRLNGRPVAVVAAHSGFALTYQFLARLARLWRFVKAVALLAALGSAFGAMD 619
 QY 253 TFASGNTLVGPVLDPLVLRDQSSSESNLMLPNPKIFGPOKPIVTPRPYSAHMVDV 312
 DB 620 GLASGAGSGPLNLPAPARLARLSQGSALWRLPVMFG-DRPVVVKNSYANNITEF 678
 QY 313 LKDIGFPEGVYPERILPLIGNIKAPQVPTITMGVGTLETLFGKGDPERPEISY 372
 DB 679 LDAIGFEGVQPYTRYLPMRRLPAPVPTSMYGVGRTPETFGVGEAFEGTPEVY 738
 QY 373 GDGQGTNVLYSLALQSLMKEEKQYLLKVKYIDGSHTSILKDEVALNEIYGEI----TS 428
 DB 739 GDGQGNMNIYSLMAAE-WSGVEQIILKVRLLPVSHVGFPSD-LALKKYVAEIQKAVSS 796
 QY 429 INSHAELGLSNLFS 442
 DB 797 IEVHAK--EKIFS 807

RESULT 3
 Q9SDR9 PRELIMINARY; PRT; 266 AA.
 AC Q9SDR9;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Lecithin-cholesterol acyl transferase (Fragment).
 GN LCAT1.
 OS Prunus dulcis (Almond) (Prunus amygdalus).
 OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
 OC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
 OX NCBI_TaxID=3755;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. BOA CASTA;
 RA Ma R.C., Oliveira M.M.;

"Molecular Cloning of A Lecithin-Cholesterol Acyl Transferase Homolog
RT Gene from Almond (Prunus dulcis)."
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF209909; AAF22841.1; -
DR InterPro: IPR003386; LACT.
DR InterPro: IPR000379; Ser_estrs_site.
DR Pfam: PF02450; LACT; 1.
KW Transferase.
FT NON TER 1 1
SQ SEQUENCE 266 AA; 29613 MW; 057533E294332F37 CRC64;

Query Match 34.0%; Score 792.5; DB 10; Length 266;
Best Local Similarity 59.7%; Pred. No. 1,1e-60;
Matches 160; Conservative 35; Mismatches 48; Indels 25; Gaps 4;

QY 181 QVGSKFLDKNLIEASNSNGKRVLLSHSLGFLVQLLNRPSPSRKKFIKHFAL 240
D 3 RVSXKFLDLDLLENASTSNGRVLVSHSLGFLALHLNRPSPSRKKFIKHFVSL 62
QY 241 SAPMGATIDEMVTFASGTLGVPLVDPLVDEORSSSNLMILPNPKIFGQKPIVTP 300
D 63 STPMGQIVDEMLTFASGTLGVPLVDPLVDEORSSSNLMILPNPKIFGKPLVTP 122
QY 301 IRPYSAHDMVDFLKDIFGPEGVYPERILPLIGNIKAPOVITCIMGVGTLETFY- 359
D 123 SATYSASDIPOFLSDIGEGGVHPEKTRVLGLMDLVAPG-----GAYVLCFWK 171
QY 360 ---GKGF-----DERPEISGDDGVNLVSLALQSLWKEKNQYLKVKIDVS 408
D 172 WCESGDGEVLWKQXVLMSSRLFLMGDDGVNVMVSLMAHESIMDKMKTLTIRIPVS 231
QY 409 HTSLKDEVALNEIYGETSINSNAELG 436
D 232 HTAILKDEGALDQIVRETSGINS--QLG 257

RESULT 4
Q8RYV9 PRELIMINARY; PRT; 363 AA.
AC Q8RYV9;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE P0614D08.29 protein.
GN P0614D08.29.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa niponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0614D08.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP004073; BAB86261.1; -
SQ SEQUENCE 363 AA; 39604 MW; 983FE2E5C67A9252 CRC64;

Query Match 28.9%; Score 673.5; DB 10; Length 363;
Best Local Similarity 42.1%; Pred. No. 4e-50;
Matches 153; Conservative 63; Mismatches 126; Indels 21; Gaps 9;

QY 80 RLMPDSSVY-----LAPFQCFARMTLHYQELDDYFNTPGVETRPVPHFGSTNSLYLN 134
D 9 RLMEVFTALQEDPALSP---CYADQLRLVYDVPVAGDYNVNCVDRVVSFGSTRGRFDD 65
QY 135 PRLKHITGYMAPLVDSLOKLGADGETLFGAPYDFRYGLAEGHPSQVSKFLDKLNL 194
D 66 PARQDVC--MERIVALEEIEEGEAGENLFGAPYDFRYAPAAPGLPSGVSGFPTSLRLRV 123
QY 195 EASNSNNGKRVLLSHSLGFLVQLLNRPSPSRKKFIKHFILSLAPMGATIDEMVTF 254

124 EASASRNGKRPVILVTHSLGFLFAMVFLDRPLPARRRIKHFVMLCGVGSSPLMMPL 183
QY 255 ASGNLTGVP---LYDPLLVDEORSSSNLMILPNPKIFGQKIVTTPRPSAHMV 310
D 184 AAS--IPSSSLGSLVLTG-VNRSFASMSPLSPAVYG-DFPLVITRAKNYSADMP 238
QY 311 DFLKDIGFPEG-VYPERILPLIGNIKAPOVITCIMGVGTLETFYFGKDFDERPE 369
D 239 EFLSAGFSDDVALYRRALPVLIDFRAPLVLSINGVGVPYDKLVFMDGNFSAKQ 298
QY 370 ISYGDGTVNLVSLALQSLWKEKNQ-YLKVKIDVSHTSILKDEVALNEIYGETS 428
D 299 VVNGGDDQINDIYVLAERLRLGDDPDGFRFSILIPMTHTKGVSDLEALKRVSEILG 358
QY 429 INS 431
D 359 ANS 361

RESULT 5
Q8VEB4 PRELIMINARY; PRT; 412 AA.
AC Q8VEB4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Similar to LCAT-like lysophospholipase (lysosomal phospholipase A2).
DE A2).
GN C87498 OR LPLA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=KIDNEY;
RA Hiraoka M., Abe A., Shayanman J.A.;
RT "Cloning and characterization of a lysosomal phospholipase A2, 1-0-acylceramide synthase.";
RL J. Biol. Chem. 0:0-0(2002).
DR EMBL: BC019373; AAL19373.1; -
DR EMBL: AF468958; AAL78651.1; -
DR MGI: 2142884; C87498.
DR InterPro: IPR003386; LACT.
DR Pfam: PF02450; LACT; 1.
SQ SEQUENCE 412 AA; 47307 MW; 6544369410C46C9A CRC64;

Query Match 26.2%; Score 612; DB 11; Length 412;
Best Local Similarity 34.2%; Pred. No. 1,1e-44;
Matches 139; Conservative 83; Mismatches 161; Indels 24; Gaps 10;

QY 16 LTVTVVVMVSLCTCGASNDPLILIPGNGNQLAARTNQKRPST--FICESVYPLIK 73
D 17 LLLPLFLMLMLADLTLPQORHPPVVLVPGDLGNQLEAKID--KPKVVHYLCS-----K 67
QY 74 KKNKGFRLMPDSSVYLAFTQCFARMTLHYQELDDYFNTPGVETRPVPHFGSTNSLYL 133
D 68 KIDSFTTLMNLLELLPVIIDGWIDINIRLVNRTSRATQPDGVVAVPGFGETPSMEFL 127
QY 134 NPLKHITGYMAPLVDSLOKLGADGETLFGAPYDFRYGLAEGHPSQVSKFLDKLNL 193
D 128 DPSKNVGSYFTWVESLVGWCYTGEDYKGRAPYMKRA-----PNNGPFLFLA-LKEM 180
QY 194 IEASNSNNGKRVLLSHSLGFLVQLLNRPSPSRKKFIKHFILSLAPMGATIDEMVTF 253
D 181 IEMVQYGG--FVVLVAHSMGNVYMLYFLQRPQVWKDKYIIAHSFGAPMGVAKTLRV 239
QY 254 PASGNTLGVPLVDPLVDEORSSSNLMILPNPKIFGQKPIVITTPRPSAHMVDPL 313

```
Db 240 LASDNMRIPVIGPKTRKREQORSASVSTSWLLPYNHTWSEKVFYPTNTNLTDRYHRNF 299
QY 314 KDIFPEGVYPTETRIPLINIKAKQVPTICMGVGTLETLEFGKGDDE-PEISY 372
Db 300 RDIPEGDMFMRQD-TEGLVEMKTPPGVELHCLGTGTPPTNSRY--ESPDPDKICF 356
QY 373 GGDGTVNLVSLALQSLMKKEKNQYLKVKIDGVSHTSLKDEVAL 419
Db 357 GGDGTVNLVSLALQSQ-MQSRQEHVSLAQELPGSEHIEMLANNTTL 402

RESULT 6
QY283 PRELIMINARY; PRT; 412 AA.
ID QY283
AC QY283;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE LCAT-like lysophospholipase (LPL).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99194552; Pubmed=10092508;
RA Tanigawa Y., Shibata S., Kita S., Horikoshi K., Shirafuji H.,
RA Sumino Y., Fujino M.;
RT "Cloning and expression of a novel lysophospholipase which
RT structurally resembles lecithin cholesterol acyltransferase.";
RL Biochem. Biophys. Res. Commun. 257:50-56(1999).
DR EMBL: AB017494; BAA76877.1; -.
DR InterPro: IPR003386; LACT.
DR Pfam: PF02450; LACT; 1.
SQ SEQUENCE 412 AA; 46657 MW; 1FEA8A5783AF050A CRC64;

Query Match 26.1% Score 610; DB 4; Length 412;
Best Local Similarity 34.8%; Pred. No. 1.6e-44;
Matches 141; Conservative 80; Mismatches 156; Indels 28; Gaps 12;

QY 21 VVWVLSLLC--TCGASNDPLILIPGNGNQLERLTNQYKRPST--FICESMYPLIKKN 76
Db 20 LLLMLLADPALPAGKRPVVLVPGDLGNLEKLD---KPYVHYLCS-----KTE 70
QY 77 GWERLWDSVYLAFTQCAERMTLHYHOELDYFNTPGVETRVPHFGSTNSLLYINPR 136
Db 71 SYFTIMNLELLPVIIDCMIDNIRLVNKTSRATQPPDGDVAVPGGKFFSLFDIPS 130
QY 137 LKHITGNAPLVDSLQKLGADGETLFGAPIDFVYGLAEGHPSQVSKFLDKNLIEE 196
Db 131 KSSVGSYFHTWVESLVGMGYTRGEDVRCAPDWMRA-----PNENGPFILA-LREIMEE 183
QY 197 ASNSNGKPYILSHSLGGLVLOLLNRNPPSWMKKFKIKHIALSAPMGGAIDEMYFAS 256
Db 184 MYQLYGG-PVYLAHNSGNNMTLYFLQROPQAMKDKITRAVFSGAGWGYAKTLRVLAS 242
QY 257 GNTLGVPLVPLVLRDORSSESNTMLLPNKIFGPKPIVTPPIRYSADWVFLDKI 316
Db 243 GDNMRIVIVIGPKTRKREQORSASVSTSWLLPYNHTWSEKVFYPTNTNLTDRYHRNF 302
QY 317 GFPEEG-VYPTETRIPLINIKAKQVPTICMGVGTLETLEFGKGDDE-PEISYGD 374
Db 303 GFEDGWMLRQDTE--GLVEATMPGVOHLCLYGVPTPSFY--ESFPDRDKICFGD 358
QY 375 GGDGTVNLVSLALQSLMKKEKNQYLKVKIDGVSHTSLKDEVAL 419
Db 359 GGDGTVNLVSLALQSQ-MQSRQEHVSLAQELPGSEHIEMLANNTTL 402

RESULT 7
QY283 PRELIMINARY; PRT; 407 AA.
ID QY283
AC QY283;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Lysosomal phospholipase A2.
GN LPLA2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX Hiraoka M., Abe A., Shayanman J.A.;
RA "Cloning and characterization of a lysosomal phospholipase A2, 1-O-
RT acylceramide synthase.";
RL J. Biol. Chem. 0:0-0(2002).
DR EMBL: AY072914; AAL65270.1; -.
DR InterPro: IPR003386; LACT.
DR Pfam: PF02450; LACT; 1.
SQ SEQUENCE 407 AA; 46062 MW; CEASAB747E7EB8FC CRC64;

Query Match 25.1% Score 586; DB 6; Length 407;
Best Local Similarity 35.1%; Pred. No. 1.9e-42;
Matches 137; Conservative 72; Mismatches 157; Indels 24; Gaps 10;

QY 33 ASNDPLILIPGNGNQLERLTNQYKRPST--FICESMYPLIKKNCFRLMFDVSYLA 90
Db 29 AGSRPVLVPGDMGNLEKLD---KPYVHYLCS-----KRTDYFTLMNLELLP 79
QY 91 PFTQCFERMTLHYHOELDYFNTPGVETRVPHFGSTNSLLYINPRKHITGNAPLVDS 150
Db 80 VIIDCMIDNVRILYNQTSHTQPPDGDVAVPGGDFESMEFLDPSKSSVGSYLHTWVES 139
QY 151 LQKLGVDGETLFGAPIDFVYGLAEGHPSQVSKFLDKNLIEESNSNNGKPVLLS 210
Db 140 LVSWGTERGKDYKAPDWMRA-----PNDNGYFLA-LKMTIEEYQLYGG-PVYLA 191
QY 211 HSLGGLVLOLLNRNPPSWMKKFKIKHIALSAPMGGAIDEMYFASGNTLGVPLVPLV 270
Db 192 HSMGNMYMLVFLQHPQDMKDYIRAVVAGPMPGVPTKTLRVLASDNMRIPVIRSLKI 251
QY 271 RDEORSSESNTMLLPNKIRGPKPIVTPPIRYSADWVFLDKIDFPEGVYPTETRI 330
Db 252 RAQORSASVSTWLLPYSTWSPQVFVTRPKANTLQDYQFQDQIFGKDG-MSMRQDTE 310
QY 331 PLIGINAKQVPTICMGVGTLETLEFGKGDDE-PEISYGGDGTNLVSLALQSL 389
Db 311 GLVEATVPPGVRLHCLGTGVPPTPSFY--ESFPDRDKIHYGTGTVNLQSLALCHT 368
QY 390 LMKKEKNQYLKVKIDGVSHTSLKDEVAL 419
Db 369 -MRGLQKQEVSLQALPGNEHIALNANTTL 397

RESULT 8
QY2168 PRELIMINARY; PRT; 421 AA.
ID QY2168
AC QY2168;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE BCDNA.GH02384.
GN BCDNA.GH02384 OR CG18858.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX Rudin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Agbayanl A., Arcaina T.T., Baxter E., Blazek R.G., Butenhoff C.,
```

```
RA  Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA  Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA  Houston K.A., Hummasti S.R., Kim E., Li P., Mostrefi M., Pacleb J.M.,
RA  Park S., Sequiera A., Sethi H., Snir E., Svitslks R.R., Weinburg T.,
RA  Celinker S.E.;
RT  "Full length Drosophila melanogaster cDNA sequence.";
RL  Submitted (APR-1999) to the EMBL/Genbank/DBJ databases.
DR  EMBL: AF145599; AAD38574.1; -
DR  FlyBase: FBgn0027612; BCdNA:GH02384.
DR  InterPro: IPR003386; LACT.
DR  Pfam: PF02450; LACT; 1.
SQ  SEQUENCE 421 AA; 48125 MW; C7F65A0F7AA32DFE CRC64;

Query Match      24.4%; Score 570; DB 5; Length 421;
Best Local Similarity 31.8%; Pred. No. 5e-41;
Matches 124; Conservative 73; Mismatches 171; Indels 22; Gaps 7;

QY  36 LDPLILIPNGNGQLERLRTNQYKPSSTFICESWYPLIKKKNGFRLMFDSSVILAFPTQC 95
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  44 LSPVIFVPGDGGSGMDARLNKRNSP-----YLICQKTHDMYINMLDLEQLVITPVYC 95
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY  96 FAERMTLHYHQELDDYFNTPGVETRVPHGSGTNSLLYLNPRLKHITGYAPLVDSLOKLG 155
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  96 WIDNVKLYYDKYRTHTNTPGVETRIIPGNGNPEVEMIDPTKNSAGAYFKDIANELVKLG 155
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY  156 YADGETLFGAPYDFRYGLAEGHPSOYGSFKLDKLNLEASNSNGKRPVILSLSLG 215
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  156 YIRKQNIHGAPYDFRKA-----PNE-NOQFFIDLQVLQVDEYSEANQOSAVTFISHSGS 208
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY  216 LFLVLQILNRPSPSRKKFKIFKIFALSAWGAIDEMYTASGNTLVPLVPLLVDEOR 275
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  209 LMTIVFLQEQTLQWKKAKYKRMISLAGWAGSKAVKAVFAMGDDLSFALSATILAEQI 268
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY  276 SSSSNMLILPNPKIFGQPKIVITTPRPSAHDMVDFLKGDFEGEVYETR--ILPLI 333
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  269 THPSTAMLPLSPFLMKPTEVLAMTPSRNYTMAOLEEFGKDLDMYTG--WEMKQDITRYN 325
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY  334 GNIRKAPQVPTICMGVGTLETLFYGKGDFD-ERPEISYGGDGVNVLALSLQSLMK 392
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  326 RNFDPPNVELHMLXGEGIDVERLOQYKKSISGTEPKLIMGLDGTVNORSILACK-YYA 384
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY  393 EERNQYLKVKKIDGVSHTSILKDEVALNEI 422
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  385 GYSKAPINTLALQNVDLHLILSNPDVLKXI 414
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 9
Q917L9      PRELIMINARY;      PRT;      421 AA.
AC  Q917L9;
DT  01-MAR-2001 (TREMBLrel. 16, Created)
DT  01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE  CG18858 protein.
DE  01-MAR-2002 (TREMBLrel. 20, Last annotation update)
GN  BCdNA:GH02384 OR CG18858.
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC  Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
OX  NCBI_Taxid=7227;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=BERKELEY;
RA  MEDLINE=20196006; PubMed=10731132;
RA  Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA  Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA  George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA  Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA  Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA  Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA  Abril J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA  Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA  Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA  Borikova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA  Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA  Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA  de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA  Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA  Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA  Flosser C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA  Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA  Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA  Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA  Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA  Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA  Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA  Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA  Merkulov G., Misha N.V., Mobarry C., Morris J., Mostrefi A.,
RA  Mount S.M., Moy M., Murphy B., Murphy L., Muny D.M., Nelson D.L.,
RA  Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA  Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA  Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA  Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA  Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA  Svitslks R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA  Wang Z.-Y., Maestran D.A., Weinstein G.M., Weisenbach J.,
RA  Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA  Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA  Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA  Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT  "The genome sequence of Drosophila melanogaster.";
RL  Science 287:2185-2195(2000).
DR  EMBL: AE003666; AAG22446.1; -
DR  FlyBase: FBgn0027612; BCdNA:GH02384.
DR  InterPro: IPR003386; LACT.
DR  Pfam: PF02450; LACT; 1.
SQ  SEQUENCE 421 AA; 48212 MW; 7F9A299B253287DA CRC64;

Query Match      24.0%; Score 560; DB 5; Length 421;
Best Local Similarity 31.5%; Pred. No. 3.7e-40;
Matches 123; Conservative 73; Mismatches 172; Indels 22; Gaps 7;

QY  36 LDPLILIPNGNGQLERLRTNQYKPSSTFICESWYPLIKKKNGFRLMFDSSVILAFPTQC 95
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  44 LSPVIFVPGDGGSGMDARLNKRNSP-----YLICQKTHDMYINMLDLEQLVITPVYC 95
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY  96 FAERMTLHYHQELDDYFNTPGVETRVPHGSGTNSLLYLNPRLKHITGYAPLVDSLOKLG 155
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  96 WIDNVKLYYDKYRTHTNTPGVETRIIPGNGNPEVEMIDPTKNSAGAYFKDIANELVKLG 155
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY  156 YADGETLFGAPYDFRYGLAEGHPSOYGSFKLDKLNLEASNSNGKRPVILSLSLG 215
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  156 YIRKQNIHGAPYDFRKA-----PNE-NOQFFIDLQVLQVDEYSEANQOSAVTFISHSGS 208
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY  216 LFLVLQILNRPSPSRKKFKIFKIFALSAWGAIDEMYTASGNTLVPLVPLLVDEOR 275
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  209 LMTIVFLQEQTLQWKKAKYKRMISLAGWAGSKAVKAVFAMGDDLSFALSATILAEQI 268
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY  276 SSSSNMLILPNPKIFGQPKIVITTPRPSAHDMVDFLKGDFEGEVYETR--ILPLI 333
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  269 THPSTAMLPLSPFLMKPTEVLAMTPSRNYTMAOLEEFGKDLDMYTG--WEMKQDITRYN 325
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY  334 GNIRKAPQVPTICMGVGTLETLFYGKGDFD-ERPEISYGGDGVNVLALSLQSLMK 392
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  326 RNFDPPNVELHMLXGEGIDVERLOQYKKSISGTEPKLIMGLDGTVNORSILACK-YYA 384
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY  393 EERNQYLKVKKIDGVSHTSILKDEVALNEI 422
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  385 GYSKAPINTLALQNVDLHLILSNPDVLKXI 414
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 10
Q9M204      PRELIMINARY;      PRT;      440 AA.
AC  Q9M204;
DT  01-OCT-2000 (TREMBLrel. 15, Created)
```

```
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE lecithin cholesterol acyltransferase.
OS Tupaiidae (Tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.
OX NCBI_TaxID=9395;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang J., Chen B.B., Zeng W.W., Wu G., Zhang C.W., Zhang M.K., Xue H.;
RT "Cloning and sequencing of tree shrew lecithin cholesterol
acetyltransferase cDNA."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF272861; AAF78242.1; -.
DR InterPro; IPR003386; LACT.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estra_site.
DR Pfam; PF02450; LACT; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
KW Acyltransferase; Transferase.
SQ
SEQUENCE 440 AA; 50281 MW; A79867564448CAC4 CRC64;

Query Match 23.4%; Score 546; DB 6; Length 440;
Best Local Similarity 33.3%; Pred. No. 6.5e-39;
Matches 129; Conservative 74; Mismatches 162; Indels 22; Gaps 11;

QY 38 PLILPGNGNQLERLNTQKPSFFICESWPLIKKNGWFRMLFSDSVILAPFTQGA 97
D 48 PVLVPGCGNGNLEAKLD---KPDVY---NMW-CYRKEDFTYVLNLTFLPVGDCMT 100
QY 98 ERMTHYHOELDDYFNFTGVEFRVPHFGSTNSLYLNPRLKHITGYMAPLVDSLOKGYA 157
D 101 DNTRTVYNNSTSKMSNAGCVHTRVPGKTYVEVLDD--SKLACYLHTLVGNLNNNGYV 158
QY 158 DGETLFGAPYDERYGLAAGHPGVSQVSKFLDKNLIEBASNNNGKPYILSHSLGLF 217
D 159 RDEIVRAAPYDWRG-----PQO-QEYKRDALRLEE-MHATYGRKPFVLGHSLGCIH 210
QY 218 VLQLLRNPSPSWRRKKFIKFTALSPMGCAIDEMTFASGNTLGVPLVDPLVLRDQRSS 277
D 211 LHLFLLHQPSQWKDRFDIGFISLAPMGWSIKPMQVLASGDNQGIPIINSIKLKEQRIT 270
QY 278 ESNMLLPNPKIFGQKPIVPIPIRYSANDVDFLKDIGFEGYVPEYTRILPLIGNIK 336
D 271 TTSPMFMFSSSEWPEPDHFIETSPSYNTYIRYQRFVVDVHFEEGYMMQLQR--DLLAGL 328
QY 337 KAPQVPICTMGVGTLETLEFYGKG-DPDERPEISYGDGDTVNLVSLALOSLMEKRN 395
D 329 PPGVETVCLYGVGLPTPESTYIYDHPPTYPDPLVDLVLEDGDMTVATRS-MELCSOMQGRQ 387
QY 396 NQYLVKVIDGVSHTSILKDEVALNEI 422
D 388 PQVHLPLHLRIQHLNMFNSNOTLEHI 414

RESULT 11
ID Q9DDJ6 PRELIMINARY; PRT; 451 AA.
AC Q9DDJ6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE lecithin cholesterol acyltransferase.
OS Anas platyrhynchos (Domestic duck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
OX NCBI_TaxID=8639;
RN [1]
RP SEQUENCE FROM N.A.
RA Zeng W.W., Zhang J., Chen B.S., Wu G., Zhang W.C., Bai L., Xue H.;
RT "Cloning and sequencing of duck lecithin cholesterol acyltransferase
cDNA.";
```

```
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF324887; AAG42498.1; -.
DR InterPro; IPR003386; LACT.
DR Pfam; PF02450; LACT; 1.
KW Acyltransferase; Transferase.
SQ
SEQUENCE 451 AA; 51469 MW; A19635D7FA37F874 CRC64;

Query Match 22.9%; Score 533.5; DB 13; Length 451;
Best Local Similarity 31.8%; Pred. No. 8.3e-38;
Matches 130; Conservative 80; Mismatches 168; Indels 31; Gaps 12;

QY 38 PLILPGNGNQLERLNTQKPSFFICESWPLIKKNGWFRMLFSDSVILAPFTQGA 97
D 47 PVLVPGCGNGNLEAKLD---KPDVY---NMW-CYRKEDFTYVLNLTFLPVGDCMT 99
QY 98 ERMTHYHOELDDYFNFTGVEFRVPHFGSTNSLYLNPRLKHITGYMAPLVDSLOKGYA 157
D 100 DNTRTVYNNSTSKMSNAGCVHTRVPGKTYVEVLDD--SKLACYLHTLVGNLNNNGYV 157
QY 158 DGETLFGAPYDERYGLAAGHPGVSQVSKFLDKNLIEBASNNNGKPYILSHSLGLF 217
D 158 RDOIVRAAPYDWRG-----PQOPEYFQNLKALIEEMHDEYQOR-VFLTAHSGMNLN 209
QY 218 VLQLLRNPSPSWRRKKFIKFTALSPMGCAIDEMTFASGNTLGVPLVDPLVLRDQRSS 277
D 210 VLYFLQORQAWKDDYIGGFISLAPMGWSVAPLVLASGDNQGIPIINSNIKLRERQWT 269
QY 278 ESNMLLPNPKIFGQKPIVPIPIRYSANDVDFLKDIGFEGYVPEYTRILPLIGNIK 337
D 270 TTSPMFMFSLAMPEDHFIETSPSYNTYIRYQRFVVDVHFEEGYMMED--MKDLKGLP 328
QY 338 APQVPICTMGVGTLETLEFYGKG-DPDERPEISYGDGDTVNLVSLALOSLMEKRN 396
D 329 PPGVDTCLYGVGLPTPESTYIYDHPPTYPDPLVDLVLEDGDMTVATRS--SELCKRNRQOK 387
QY 397 QYLVKVIDGVSHTSILKDEVAL---NET-----VGETTSINSRAELG 436
D 388 QKVHTQELRGIDHLNMFNSNLTLSSINELLGSSQVAGAT--KEHGEIG 434

RESULT 12
ID Q35849 PRELIMINARY; PRT; 440 AA.
AC Q35849;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE lecithin:cholesterol acyltransferase (EC 2.3.1.43).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-SPRAGUE-DAWLEY, AND WISTAR; TISSUE=LIVER;
RC MEDLINE=97363611; PubMed=9219904;
RA Wang J., Gebre A.K., Anderson R.A., Parks J.S.;
RT "Cloning and in vitro expression of rat lecithin:cholesterol
acyltransferase."
RL Biochim. Biophys. Acta 1346:207-211(1997).
DR EMBL; U62803; AAB65771.1; -.
DR InterPro; IPR003386; LACT.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estra_site.
DR Pfam; PF02450; LACT; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
KW Acyltransferase; Transferase.
SQ
SEQUENCE 440 AA; 49882 MW; 836BB2D599F72B63 CRC64;

Query Match 22.5%; Score 526; DB 11; Length 440;
Best Local Similarity 33.1%; Pred. No. 3.6e-37;
Matches 129; Conservative 74; Mismatches 159; Indels 28; Gaps 12;
```

```
QY 38 P L I L I P G N G N O L E A R L T N O Y K R P S - - T F I C E S W Y P L I K K K N G M F R L M P D S S V I L A P F T Q C 95
D B 48 P V I L V P G C G N R L E A K L D - - - K P N V N W L C - - - - - Y R K T E D F F T I W D F N F L P I G V D C 98
QY 96 F A E R M T L H Y H O E L D Y F N T P G V E T R P H G S T N S L Y L N P R L K H T I G Y A P V D S I O K J G 155
D B 99 W I D N T R V Y V N R S S G H S N A P G V O I N P G R K T Y S E V E Y D D - - N K L A G I L H T L V O N L V N N G 156
QY 156 Y A D G E T L F G A P Y D F R Y G L A E G H P S O V S K F L K D K N L I E E A S N S N N G K P V I L L S H S I G G 215
D B 157 Y R D E T V R A A P Y D W R L A - - - - - P R Q - O D E Y Y K K L A G I V E E - M Y A A Y K P V F L I G H S L G C 208
QY 216 L E V I L O L N P R P S W R K K F I K H F I A L S A P W G A I D E M T F A S G N T L G V P L V D P L V D E R 275
D B 209 L H V L H F L R Q O S W K D H F I D G F I S L G A P W G S I K P R I L A S G D N O G I P T M S N I K L E E O R 268
QY 276 S E S N L M L P N K R I G P Q K P I Y T P I R P Y S A H D W F L K D I G F P E G Y P Y - E T R I L P L I G N I 334
D B 269 I T T S P M M P R A H W P E D H V F I S T P N F N Y T Q G D F E R F F D L H F E B G W M F L Q S R - - D L L A 326
QY 335 N I K A P O V P I T C I M G T G V G T L E T F Y G K G D F D E R P E I S - - Y G D G D T V N L V S L A L O S L M K 392
D B 327 G L P A G V E Y C L Y G V G M P A H T Y I Y D H - N F P Y K D P V A L Y E D G D D T V A T R S - T E L C G O M Q R 384
QY 393 E E K N O Y L K V K I D G V S H T S I L K D E V A L N E I 422
D B 385 G R O S A V H L P M N G T D H L M V F S N K T L E H I 414

RESULT 13
Q9CM47 PRELIMINARY: PRT: 444 AA.
AC Q9CM47:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE lecithin cholesterol acyltransferase (Fragment).
GN LCAT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LIVER;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinaagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
DR EMBL: AK004914; BAB23665.1; -
DR MGI: 96755; Lcat.
DR InterPro: IPR003386; LACT.
DR InterPro: IPR000734; Lipase.
DR InterPro: IPR000379; Ser_estrs_site.
DR Pfam: PF02450; LACT. 1.
DR PROSITE: PS00120; LIPASE_SER; 1.
```

```
FT NON_TER 1 1
SQ SEQUENCE 444 AA; 50535 MW; 7936970447639A63 CRC64;
Query Match 22.2%; Score 518; DB 11; Length 444;
Best local similarity 33.2%; Pred. No. 1.8e-36;
Matches 129; Conservative 71; Mismatches 164; Indels 24; Gaps 12;
QY 38 P L I L I P G N G N O L E A R L T N O Y K R P S - - T F I C E S W Y P L I K K K N G M F R L M P D S S V I L A P F T Q C 97
D B 54 P V I L V P G C G N R L E A K L D - - - K P D V V - - - N M - C Y R K T E D F F T I W D F N F L P I G V D C W I 106
QY 98 E R M T L H Y H O E L D Y F N T P G V E T R P H G S T N S L Y L N P R L K H T I G Y A P V D S I O K J G Y A 157
D B 107 D N T R I V N H S S G R V S N A P G V O I N P G R K T Y S E V E Y D D - - N K L A G I L H T L V O N L V N N G T V 164
QY 158 D E T T L F G A P Y D F R Y G L A E G H P S O V S K F L K D K N L I E E A S N S N N G K P V I L L S H S I G G L F 217
D B 165 R D E T V R A A P Y D W R L A - - - - - P R Q - O D E Y Y K K L A G I V E E - M Y A A Y K P V F L I G H S L G C L H 216
QY 218 V I Q L N R N P P S W R K K F I K H F I A L S A P W G A I D E M T F A S G N T L G V P L V D P L V D E R O R S 277
D B 217 V H F I L R Q O S W K D H F I D G F I S L G A P W G S I K A M R I L A S G D N O G I P I L S N I K L E E O R Y T 276
QY 278 E S N L M L P N K R I G P Q K P I Y T P I R P Y S A H D W F L K D I G F P E G Y P Y - E T R I L P L I G N I 336
D B 277 T T S P M L P A P H W P E D H V F I S T P N F N Y T Q D F E R F F T D L H F E B G W M F L Q S R - - D L L E R L 334
QY 337 K A P O V P I T C I M G T G V G T L E T F Y G K G D F D E R P E I S - - Y G D G D T V N L V S L A L O S L M K E E 394
D B 335 P A P G V E Y C L Y G V G R P T P H T Y I Y D H - N F P Y K D P V A L Y E D G D D T V A T R S - T E L C G O M Q R 392
QY 395 K N O Y L K V K I D G V S H T S I L K D E V A L N E I 422
D B 393 Q S Q P V H L P M N E T D H L M V F S N K T L E H I 420

RESULT 14
Q21515 PRELIMINARY: PRT: 417 AA.
AC Q21515:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE M05B5.4 protein.
GN M05B5.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gardner A.E.;
RN Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RT Science 282:2012-2018(1998).
RL EMBL: 271265; CA959833.1; -
DR InterPro: IPR002453; Beta_tubulin.
DR InterPro: IPR003386; LACT.
DR InterPro: IPR000379; Ser_estrs_site.
DR Pfam: PF02450; LACT. 1.
DR PROSITE: PS00228; TUBULIN_B_AUTOREG; 1.
SQ SEQUENCE 417 AA; 46839 MW; EBBB0F403F9053B CRC64;
Query Match 22.0%; Score 512.5; DB 5; Length 417;
Best local similarity 32.8%; Pred. No. 4.9e-36;
Matches 129; Conservative 79; Mismatches 160; Indels 25; Gaps 12;
QY 38 P L I L I P G N G N O L E A R L T N O Y K R P S - - T F I C E S W Y P L I K K K N G M F R L M P D S S V I L A P F T Q C 95
```

```

Db 35 PVLIVPGGGGQLESNLG--KPSVHYVCS-----KOTADYFDLMLNLQLETPPLVIDC 86
QY 96 FAERTLLHYHOELDDYFNTPGVETRVPHFGSTNSLLYLNPRLKHITGVMAPLVDSIQKLG 155
Db 87 WADNNQVLVFNNTTGLSEMMPGVDIVAGFGATESVEMLDKSAASGRFFPDIVDSMWSG 146
QY 156 YADGETLGPARYDFRYGLAEGHPSQVSKFLDKNLIEEASNSNNGKPVLLSHSLGG 215
Db 147 YRRGNVIGAPFDMRKS-----PNEL-NDYLIQLKSLIETTYRYMNDNQIKYLVHSGN 199
QY 216 LFVLQLLNRN-PPSWRRKFIFKFIILSAPWGAIDEMTFASGNTLV--PLVDDLARD 272
Db 200 PLSLFLNNYVDQAKDKYISSFVSLAPWAGSMQIVRLFASGYNMNYRYVILPPSSLRA 259
QY 273 EORSESNLMLPNKIGPOKPIVTPIRPSAHDMDFLKIGFPEGVYPERILPL 332
Db 260 MORSFSSSAFLFPPSVAMKPHIELATTDKNTYGNIKFEPDINMYMGWEQYK-QAARL 318
QY 333 IGNIKAPVPTICMGVGTLETLEFGKGFDERPEISY-GDGDGTVNLVSLALQSLW 391
Db 319 NGNLSSPGVPYHCIGVTPPEKFSMAQYFPDYPTPEFGDGDGTYNKKSATVCTN-W 377
QY 392 KEENQY-LKYVKIDGVSHTSILKDEVALNEIV 423
Db 378 IGNNNGYKTVHEVQADHMAILKHPNAI-ELV 409

```

RESULT 15

```

09UG04
AC 09UG04; PRELIMINARY; PRT; 272 AA.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 31.0 kDa protein (unknown) (Protein for MCC:12657).
CN DKFZP564A0122.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL110209; CAB53675.1; -.
DR EMBL; BC011640; AAH11640.1; -.
DR InterPro: IPR003386; LACT.
DR Pfam; PF02450; LACT; 1.
KW Hypothetical protein.
SQ
SEQUENCE 272 AA; 31016 MW; ACCC5E1680D7A720 CRC64;

```

Query Match 19.6%; Score 457.5; DB 4; Length 272;
 Best Local Similarity 37.1%; Pred. No. 1.6e-31;
 Matches 102; Conservative 54; Mismatches 104; Indels 15; Gaps 8;

```

QY 147 LVDSIQKLGYADGETLFGAPYDFRYGLAEGHPSQVSKFLDKNLIEEASNSNNGKPV 206
Db 1 MVESLVGNGYRGEDVGRGAPYDMRRA-----PNENGPIFLA-LREMIEMQLYG-PV 52
QY 207 ILLSHSLUGLFLVQLNLNPPSMRKKFKHPIALSAPWGAIIDEMTYFASGNTLVGLVD 266
Db 53 VLVASHMGMNMTLVFLQROPQAMKDKYIRAFVSLGAPGWGAKTLRLVLASGDNNRIPYIG 112
QY 267 PLVVRDEORSESNLMLPNKIGPOKPIVTPIRPSAHDMDFLKIGFPEG-VYPY 325
Db 113 PLKIREOORSASVSTWLLPYNTWSPEKVFVOTPTINTLRDYRKRFODIGFEDGWLMO 172

```

```

QY 326 ETRILPLIGNIKAPVPTICMGVGTLETLEFGKGFDER-PEISYGDGDGTVNLVSL 384
Db 173 DTE--GLVEATMPPGVQLHCYGTGVPPTPDSFY--ESFPDRDPKICPGDGDGTVNLKSA 228
QY 385 LALQSLMKEENQYLYKVKIDGVSHTSILKDEVAL 419
Db 229 LQCOA-WQSRQEHQVLLQELPGESEHIELMANATTL 262

```

Search completed: February 11, 2003, 19:15:18
 Job time : 37 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 15, 2003, 13:17:02 : Search time 2485 seconds

(without alignments)
5188.146 Million cell updates/sec

Title: US-09-857-612a-14
Perfect score: 2333
Sequence: 1 MKKEDEGLKIEVALTVTV.....GETTSINSHALGLSLNLFSG 443

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV=rlh
-Q/c9q2_1/USFTO.spool/US09857612/r/unalc_11022003_111818_29825/app-query.fasta_1.583
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human4.0.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09857612.ecgn_1.1.1616.etrnat_11022003_111818_29825 -NCPD=6 -ICPD=3
-WARR=TIMOUT=30 -THREDOERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV=TIMEDOUT=120
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1380.5	59.2	1299	6	AX090309 Sequence
2	1380.5	59.2	1299	8	AY133614 Arabidops
3	1380.5	59.2	1457	8	AF67326 Arabidops
4	1380.5	59.2	1498	8	AY087433 Arabidops
5	1310	56.2	1562	6	AX037586 Sequence
6	1310	56.2	1562	6	AX037605 Sequence
7	1310	56.2	99690	8	AC004557 Genomic s
8	1174	50.3	134188	2	AC120983 Oryza sat
9	900	38.6	833	8	AF209909 prunus du
10	696.5	29.9	123620	8	AC103891 Oryza sat
11	675	28.9	139534	2	AP005640 Oryza sat
12	655.5	28.1	147548	8	AP004073 Oryza sat
13	612	26.2	1236	6	E26773 Novel prote
14	612	26.2	1239	10	AF468958 Novel prote
15	612	26.2	2721	10	AF468958 Mus muscu
16	610	26.1	1236	6	E26771 Novel prote
17	610	26.1	1400	9	AB017494 Homo sapi
18	610	26.1	2681	9	AK001705 Homo sapi
19	608	26.1	1137	6	E26766 Novel prote
20	608	26.1	1146	6	E26769 Novel prote
21	606	26.0	1137	6	E26768 Novel prote
22	601	25.8	2722	9	HSMB00863 Novel prote
23	596	25.5	1332	6	E26772 Novel prote
24	594	25.5	1233	6	E26767 Novel prote
25	594	25.5	1242	6	E26770 Novel prote
26	586	25.1	1224	4	AY072914 Bos tauru
27	570	24.4	1820	3	AF145599 Drosophila
28	554	23.7	1155	6	AX090327 Sequence
29	546	23.4	1338	4	AF272861 Tupai a g1
30	544	23.3	1341	4	RABLCAT
31	541	23.2	1307	9	HUMLCATGA
32	541	23.2	1355	6	E01185 DNA encodin
33	541	23.2	1360	9	BC014781 Homo sapi
34	541	23.2	1744	6	AX077318 Sequence
35	541	23.2	1744	6	AX409598 Sequence
36	541	23.2	1744	6	HUMLCAT
37	541	23.2	4308	6	AX077320 Sequence
38	537	23.0	1343	9	BABLCAT
39	536	23.0	1297	9	HUMLCATGA
40	536	23.0	1447	9	HSLCAT1
41	533.5	22.9	1953	5	AF324887 Anas plat
42	526	22.5	1346	10	RND62803 Rattus norv
43	518	22.2	1354	10	BC028861 Mus muscu
44	517.5	22.2	128210	2	AP005643 Oryza sat
45	516	22.1	1341	10	MUSLCATX J05154 Murine cho1

RESULT 1

ALIGNMENTS

[illegible]

QY	199	AanserAsnAnglVlysprOvalIleleuSerHisSerLeuGlyLeuPheVal	218
QY	199		
Db	586	AGGAGCAAGAAAGAAAGCCAGTACTCTCTCCATAGCTTGGAGAGACTTTTGGTC	645
QY	219	LeuGlnLeuLeuAsnArgAsnProProSerTPAcGlyLysPheIleLysHisPheIle	238
QY	219		
Db	646	CTTCATTTTCTCAACCGTACCAACCCCTTCATGGCGCCGGAATACATCAAAACCTTTGTT	705
QY	239	AlaLeuSerAlaProTrrpGlyLysAlaIleAspGluMetTyrThrPheAlaSerGlyAsn	258
QY	239		
Db	706	GCACCTCGCTGCCCATGGGGTGGAGACATCTCTCAGATGAGACATTTTGGCTTGGCAAC	765
QY	259	ThrLeuGlyValProLeuValIAspProLeuLeuValArgAspGluGlnArgSerSerGlu	278
QY	259		
Db	766	ACACTCGGTGTCTTTATTACTTACCTTTGTCTGGTACAGCGCATCAGAGGACCTCCAG	825
QY	279	SerAsnLeuTrpLeuLeuProAsnProLysIlePheGlyProGln---LysProIleVal	297
QY	279		
Db	826	ACTAAACCATGGCTACTCTTCATCTACCAAAAGCTTTACAGACAGAACTAAACCGCTTGTG	885
QY	298	IleThrProIleArgProTrrSerAlaHisAspMetValAspPheLeuLysAspIleGly	317
QY	298		
Db	886	GTAACTCCCGACGTTAACTACACACACTACAGAGATGCTGTTTGGTCAGACATTTGGA	945
QY	318	PheProGluGlyValTyrProTrrGlyLupThrArgIleLeuProLeuIleGlyAsnIleLys	337
QY	318		
Db	946	TTCTCTACAAAGAGGTGTGCTCTTACAAAGACAAAGATGTGCTTTTAAACAGAGAGCTGTATG	1005
QY	338	AlaProGlnValProIleThrCysIleMetGlyThrGlyValGlyThrLeuGlnThrLeu	357
QY	338		
Db	1006	ACTCCGGGAGTCCAGTCACTCATATATATGAGAGAGAGGTGATACACACCGAGCTTTTG	1065
QY	358	PheTyrGlyLysGlyAspPheAspGluArgProGluIleSerTyrGlyAspGlyAspGly	377
QY	358		
Db	1066	ATGTATGGAAAGAGAGATTCATTAAGCAGACAGATTAAGTATGAGATGGACATGGG	1125
QY	378	ThrValAsnLeuValSerLeuLeuAlaLeuGlnSerLeuTrrpLysGluGlnLysAsnGln	397
QY	378		
Db	1126	ACGGTAAATTTGGCGAGCTTACACACTTGG-----AAATCGCAT	1164
QY	398	TyrLeuLysValValLysIleAspGlyValSerHisThrSerIleLeuLysAspGluVal	417
QY	398		
Db	1165	ACCTTGAAACACCGCTAGAGATTTGATGGAGTTTGGCATCATCTATCTTAAGACAGCATC	1224
QY	418	AlaLeuAsnGluIleValAlaGlyGluIleThrSerIleAsn	430
QY	418		
Db	1225	GCACCTAAAGACATTTATGACAGATTTTCATATTATTAAT	1263
RESULT 2			
AY133614			
LOCUS	AY133614	1299 bp	mRNA linear PLN 07-AUG-2002
DEFINITION	Arabidopsis thaliana	Al1927480/F17L21_28	mRNA, complete cds.
ACCESSION	AY133614		
VERSION	AY133614.1	GI:22137197	
KEYWORDS	FLI.CDNA.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;		
	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.		
REFERENCE	1 (bases 1 to 1299)		
AUTHORS	Chouk,R., Chen,H., Kim,C.J., Shim,P., Banh,J., Bowser,L.,		
	Carinciol,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashi,Izaki,Y.,		
	Ishida,J., Jones,T., Kamlay,A., Karlin-Neumann,G., Kawai,J.,		
	Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,		
	Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,		
	Southwick,A., Tang,C.C., Tortum,M., Wu,H.C., Yamada,K.,		
	Yamanura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.		
	and Ecker,J.R.		
TITLE	Arabidopsis ORF clones		
JOURNAL	unpublished		
REFERENCE	2 (bases 1 to 1299)		
AUTHORS	Chouk,R., Chen,H., Kim,C.J., Shinn,P., Banh,J., Bowser,L.,		

Db 1165 ACCTGTAACACCGTACAGATTGATGAGTTTCGCATACATCTATACCTTAAGACGAGATC 1224
Qy 418 AAlaLeuAsnGLuIleValGluIleThrSerIleAsn 430
Db 1225 GCACCTTAAGAGATTATGAAAGCAGATTTCATTTATTTAAT 1263
RESULT 3
AF367326 1457 bp mRNA linear PLN 07-FEB-2002
LOCUS Arabidopsis thaliana Atlg27480/FL1L21_28 mRNA, complete cds.
DEFINITION AF367326
ACCESSION AF367326.1 GI:13605854
KEYWORDS FL1_CDNA.
ORGANISM Arabidopsis thaliana.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1457)
REFERENCE
AUTHORS Cheuk, R., Chen, H., Kim, C.-J., Meyers, M.C., Shinn, P., Banh, J.,
Bower, L., Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jones, T., Kamlya, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M.,
Nguyen, M., Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T.,
Satou, M., Seki, M., Southwick, A., Toriumi, M., Yamada, K., Yu, G.,
Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 1457)
Cheuk, R., Chen, H., Kim, C.-J., Meyers, M.C., Shinn, P., Banh, J.,
Bower, L., Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jones, T., Kamlya, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M.,
Nguyen, M., Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T.,
Satou, M., Seki, M., Southwick, A., Toriumi, M., Yamada, K., Yu, G.,
Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Direct Submission
Submitted (31-MAR-2001) Salk Institute Genomic Analysis Laboratory
(Stigall), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : "RIKEN
Arabidopsis Full-Length cDNA") : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamlya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.
COMMENT
The Salk, Stanford, PGSC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H.,
Cheuk, R., Kim, C.-J., Koesema, E., Meyers, M.C., Tracy, S.E., Banh, J.,
Bower, L., Chung, M.K., Goldsmith, A.D., Jones, T., Karlin-Neumann, G.,
Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,
Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Southwick, A.,
Tang, C.C., Toriumi, M., Yamada, K., Yu, G., Davis, R.W., Theologis, A.,
and Ecker, J.R.
Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.
FEATURES
Source
1. 1457
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL09-12-E12(R12274)"
/note="ecotype: Columbia"
1. 93
94. 1392
/note="unknown protein"
/codon_start=1
/product="Atlg27480/FL1L21_28"
/protein_id="AAK32913.1"

/db_xref="GI:13605855"
/translation="MKRTISSHYSVIALIVVVTMTSMQVAGSNVYPLILVPGNGNQ
LEVRIDREYKPSWSSWMLYPIHKKSGSEGRFLFEDAIVLLSPFRPSDMMLYDP
DLDDYONAPGVQTRVPHGSTRSLYLDPRIDRLDYSMEHLKALEKCGVYNQOTIL
GAPDFRYGLAAGHSPVASOPLODLKOLVKEKTSYENEGKPVLLSHLGLFLYLF
LNRTTSPSRKKRYIKHFLVLAAPMGCTISQMTKTPASGNTLSPVLPNPLVRHRTSES
NOMLIPSTKVFHDKRPLVTPVQNTYATENDREFADIGESQGVPTKTRVLPTEEL
MTPGVPTVICIGRGVDPEVLMYGGGDKQPEIKRGDGTVMNLASLALKVDSLNT
VEIDGVSHSTSLKDELALKEIMKOIISIINVELAVNAYNE"
3'UTR
BASE COUNT 405 a 318 c 336 g 398 t
ORIGIN
Alignment Scores:
Pred. No.: 1,19e-111 Length: 1457
Score: 1380.50 Matches: 264
Percent Similarity: 75.29% Conservative: 62
Best Local Similarity: 60.97% Mismatches: 92
Query Match: 59.17% Indels: 15
Gaps: 5
US-09-857-612A-14 (1-443) x AF367326 (1-1457)
Qy 1 MetLysLysGluGluGluGluGluLysIleGluValAlaThrLeuThrValThrVal 20
Db 94 ATGAAAAA-----ATATCTTCACATTATTCGGTAGTCATACGATAC 138
Qy 21 ValValAlaMetLeuSerLeuLeuSerLeuSerLeuSerLeuSerLeuSerLeuSerLeu 40
Db 139 GTTGTGTCATACATGACCCGATGTCTCAAGCTGGGGAGACAGCTGACCTTGATT 198
Qy 41 LeuIleProGlyAsnGlyGlyAsnGluGluAlaArgLeuThrAsnGluThrLysPro 60
Db 199 CCGGTTCCAGCAAGAGGAGGTAAACGAGTACGCGTGCAGACAGAAATACAAAGCA 258
Qy 61 SerThrPheIleCysGluSerTrp---TyrProLeuIleLysLysLysAsnGlyTrpPhe 79
Db 259 ACTAGTGTCTGTGATGACGAGCTGTTATATCGATTCAATGAAGAGTGGTAGGTTT 318
Qy 80 ArgLeuTrpPheAspSerSerValIleLeuAlaProPheThrGlnCysPheAlaGluArg 99
Db 319 AGCGATATGTTGATGACGACAGATGTTATTTGTTCCCTTACCAGAGTGTTCAGCATGCA 378
Qy 100 MetThrLeuHisTyrHisGluLeuLysAspTyrTrpAsnThrProGlyValGluThr 119
Db 379 ATGATGTTGATGATGACCCGATGTTGGATGATACCAAAATGCTCGTGGTCCAAAC 438
Qy 120 ArgValProHisPheGlySerThrAsnSerLeuLeuTyrLeuAsnProArgLeuLysHis 139
Db 439 CCGGTTCTCATATTGCGTTCGACCAAAATCATCTTATACCTGACCCGCTCCGAGAT 498
Qy 140 IleThrGlyTyrMetAlaProLeuValAspSerLeuGln---LysLeuGlyTyrAlaAsp 158
Db 499 GCCACATCTTACATGAGCAATTTGGTGAAGCTCTAGAGAAAAATGCGGGATGTAAAC 558
Qy 159 GlyLutThrLeuPheGlyAlaProTyrAspPheArgTyrGlyLeuAlaAlaGluLysHis 178
Db 559 GACCAAAACCATCTGAGGCTCATATGATTTTACAGTACGGCGCTGCTGGGGCCAC 618
Qy 179 ProSerGlnValGlySerLysPheLeuLysAspLeuLysAsnLeuIleGluGluLysSer 198
Db 619 CCGTCCCGTGTAGCCTCAGCTCAGTTCTTACAAACCTCAATTTGGTGAATAAACTAGC 678
Qy 199 AsnSerAsnAsnGlyLysProValIleLeuLeuSerHisSerLeuGlyLysLeuPheVal 218
Db 679 AGCGAGACGAAGAAACCACTGATCTCTCTCCATACCTTAGGGAGCTTTTCTGC 738
Qy 219 LeuGlnLeuLeuAsnArgAsnProProSerTrpArgLysLysPheIleLysHisPheIle 238
Db 739 CTCATTTCTTCACACCGTACACCCCTTCATGCGGCCGCAAGTACATCAACACTTGT 798
Qy 239 AlaLeuSerAlaProTyrGlyLysAlaIleAspGluMetTyrThrPheAlaSerGlyAsn 258

Db 799 GCACTCCGTCGGCAGTGGGGTGGGAGCAGATCTCTCAGATGAAGACATTTGCTTCTGGCAAC 858
QY 259 ThrleuGlValProleuValAspProleuValAlaArgAspGluArgSerSerGlu 278
Db 859 ACACCTGGGTGTCCTTAGTTAACCTTCTGTCTGTCTGACAGCGCATTCAGAGACTCCGAG 918
QY 279 SerAsnLeuTrpLeuLeuProAsnProLysIlePheGlyProGln---LysProIleVal 297
Db 919 AGTACCAAAAGGCTACTTCATCTACCAAAAGTGTTCACGACAGACTAAACCCCTGTC 978
QY 298 IleThrProIleArgProTySerAlaHisAspMetValAspPheLeuLysAspIleGly 317
Db 979 GTAACCTCCAGGTTAATCTACACAGCTTACAGATGATCGCTTTTTCAGACATCTGA 1038
QY 318 PheProGluValValTyProTyArgIleThrArgIleLeuProleuIleGlyAsnIleLys 337
Db 1039 TTCTCAAGAAAGAGTTGTGCTTACAAAGAGAGTGTTCCTTTTACAGAGAGCTGATG 1098
QY 338 AlaProGluValProIleThrCysIleMetGlyThrGlyValGlyThrLeuGluThrLeu 357
Db 1099 ACCTCCGGAGTGGCCAGTCTGATATATGGAGAGAGAGTGTATACACCGAGCTTTTG 1158
QY 358 PheTyGlyLysGlyAspPheAspGluArgProGluIleSerTyGlyAspGlyAspGly 377
Db 1159 ATGATATGAAAGAGAGATTCGATACCAACAGAGATTAAGTATGAGATGAGATGGG 1218
QY 378 ThrValAsnLeuValSerLeuLeuAlaLeuGlnSerLeuTrpLysGluLysAsnGln 397
Db 1219 ACGGTAAATTTGGCGAGCTTACGAGCTTTG-----AAAGTCGAT 1257
QY 398 TyrLeuValValIleValIleAspGlyValSerHisThrSerIleLeuLysAspGluVal 417
Db 1258 AGCTTGAAACACCGTAGAGATTTGATGGAGGATTTCCGATACATCTTAAGACGAGATC 1317
QY 418 AlaLeuAsnGluIleValGlyGluIleThrSerIleAsn 430
Db 1318 GCACCTAAAGAGATTATGAGCAGATTTCATATTATAT 1356
RESULT 4
AY087433 1498 bp mRNA linear PLN 26-JUN-2002
LOCUS Arabidopsis thaliana clone 35408 mRNA, complete sequence.
DEFINITION AY087433
VERSION AY087433.1 GI:21406157
KEYWORDS FLI_CDN.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 1498)
Haaß,B.O., Volkovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. (2002) In press
2 (bases 1 to 1498)
Broyer,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
TITLE JOURNAL
REFERENCE Full-length cDNA from Arabidopsis thaliana
AUTHORS Unpublished
3 (bases 1 to 1498)
Broyer,Y., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
TITLE JOURNAL
COMMENT Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
Malibu, CA 90265, USA
This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unsplliced introns and spliced exons; one

percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
sequence is considered to be 3'-truncated if it lacks the
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the Ws or Luer ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
GenSet carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.

FEATURES
source
Location/Qualifiers
1..1498
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="35408"
82..1380
/product="unknown"
/codon_start=1
/protein_id="AAM64980.1"
/db_xref="GI:21593031"

CDS

1
/translation="MKRISHSVVAVLIVVTMTSCQAVGSNYPILIPGNGNO
LEVRLDREKPPSSVWCSSWLYPIHKKSGGFRLMFPDAVLLSPFCFSDRMLYDP
DLDDYONAPGVQTRVPHFGSKSLYLDELRLDPTASEMHLVYALKGQYVNDOTIL
GAPYDFRYGLAASGHSRVASOPLODLKQVETSSENEGKPVILSLHSGLEFLAH
LNRTTPSMRRKTKYIKHVFALAAPMGSTISOKTASGNTGSPVLPVLPVRRHRTSES
NOMRLSTKTHDRTRKPIVYTPQVNTAYTADMPDPAIDIGSGCVGPKYKRVLPDTEEL
MTFGVPTCIYGGVDTPPEYLMKGGKGFDPKPTKGTMDGDTGNLNLASLAKVDSINT
VEIDGSHSILDELAKIMQIISINELANVANVE"

BASE COUNT 425 a 322 c 343 g 408 t
ORIGIN

Alignment Scores:
Pred. No.: 1,23e-111 Length: 1498
Score: 1380.50 Matches: 264
Percent Similarity: 75.29% Conservative: 62
Best Local Similarity: 60.97% Mismatches: 92
Query Match: 59.17% Indels: 15
DB: 8 gaps: 5

US-09-857-612a-14 (1-443) x AY087433 (1-1498)

QY 1 MetLysGluGluGluGluGluGlyLeuLysIleGluValAlaThrLeuThrValThrVal 20
Db 82 ATGAAAAA-----ATATCTTCACATTAATTCGTAAGCATATGCATCTC 126
QY 21 ValValValMetLeuSerLeuLeuLysThrCysGlyAlaSerAsnLeuAspProleuIle 40
Db 127 GTTGTGGAGAGATGACCTGATGTCACAGCTGTGGTACCATATGTTACCTTGAT 186
QY 41 LeuLeuProGlyAsnGlyAsnGlnLeuGluAlaArgLeuThrAsnGlnTyLysPro 60
Db 187 CTGGTCCAGAAACGAGGAGTACAGCTAGAGTACGAGGAGCAGAAATACAGCA 246
QY 61 SerThrPheLeuCysGluSerTrp---TyrProLeuIleLysLysAsnGlyTrpPhe 79
Db 247 AGTAGTGTCTGTCTAGCAGCTGTATATCCATATCATAGAGAGTGGATGTT 306
QY 80 ArgLeuTrpPheAspSerSerValIleLeuAlaProPheThrGlnCysPheAlaGluArg 99
Db 307 AGGCTATGCTTCGATGACAGCTGTTATTTGCTCCCTTACACAGTCTTACGAGATCA 366
QY 100 MetThrLeuHisTyThrHisGlnGluLeuAspAspTyThrPheAsnThrProGlyValGluThr 119
Db 367 ATGATGTTGTACTATGACCTGATTTGGATTTATACAAATAGCTCCGGGTGTCAAACC 426
QY 120 ArgValProHisPheGlySerThrAsnSerLeuLeuTyLeuAsnProArgLeuLysHis 139
Db 427 CGGTTCTCTCTTTTCGGTTCGACCAATCACTTATATACCTCGACCGTCTCGAGAT 486
QY 140 IleThrGlyTyMetAlaProleuValAspSerLeuGln---LysLeuGlyTyArgAlaAsp 158
Db 487 GCCACATCTTACATGAAACATTTGTGTGAAGCTCTAGAGAAATAATCGGATATGTTAAC 546

QY	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
QY 171	AX037605	Sequence 29 from Patent WO0060095.	AX037605	AX037605.1	GI:11227019	Arabidopsis thaliana	Arabidopsis thaliana
Db 765							
QY 191							
Db 825							
QY 211							
Db 885							
QY 231							
Db 945							
QY 251							
Db 1005							
QY 271							
Db 1065							
QY 291							
Db 1125							
QY 310							
Db 1185							
QY 330							
Db 1245							
QY 350							
Db 1305							
QY 370							
Db 1365							
QY 390							
Db 1419							
QY 410							
Db 1464							
QY 430							
Db 1524							
QY 1524							

171	TyrGluLeuAlaAlaGluGlyHisProSerGlnGlySerLysPheLeuLysAspLeu	190	Lee, J., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukharzky, N.,
172		191	Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A.,
765	TACGGCTGCTGCTTGGGCGCCACCGCTCCGCTTACCTCACCAGTTCTCCACAGACCTC	824	Thaveri, A., Toriumi, M., Vaysberg, M., Yu, G., Federspiel, N. A.,
191	LysAsnLeuIleGluGluAlaSerAsnSerAsnAsnGlyLysProValIleLeuLeuSer	210	Theologis, A. and Ecker, J. R.
825	AAACAATTGGTGAAGAAAACATGACGAGAGAAAGAAAGCAAGCATCTCTCTCC	884	Genomic sequence for Arabidopsis thaliana BAC F17L21 from
211	HisSerLeuGlyGlyLeuPheValLeuGlnLeuLeuAsnArgAsnProProSerTTPArg	230	chromosome 1
885	CATACCCTGAGAGAGACTTTTCCTCCATTTCTCCACACCGTACACCCCTCATGGCCG	944	Unpublished
231	LysLysPheIleLysHisPheIleLeuSerAlaProTrpGlyAlaIleAspGlu	250	2 (bases 1 to 99690)
945	CGCAAGTACATCAACACTTTGTTGCACTCCCTGCCATGGGGTGGAGCATCTCTCAG	1004	Ecker, J. R.
251	MetTrpThrPheAlaSerGlyAsnThrLeuGlyValProLeuValAspProLeuLeuVal	270	Direct Submission
1005	ATGAAGCATTTGCTTCTGGCAACACACTCGGTCCCTTACGTAAACCTTGGCTGGTC	1064	Submitted (14-APR-1998) Arabidopsis thaliana Genome Center,
271	ArgAspGluGlnArgSerSerGluSerAsnLeuTrpLeuLeuProAsnProLysIlePhe	290	Department of Biology, University of Pennsylvania, 38th Street and
1065	AGACGGCATCGACAGACCTCCGACAGTAACCAATGCTACTTCATCTACCAAGTGT	1124	Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
291	GlyProGln---LysProIleValIleThrProIleArgProTrpYserAlaHisAspMet	309	4 (bases 1 to 99690)
1125	CACGACCAACTAAACCGCTTGTGTAACTCCCAAGTAAACACAGCTTACGAGATG	1184	Ecker, J. R.
310	ValAspPheLeuLysAspIleGlyPheProGluGlyValTyrrProGlyTrpArgIle	329	Direct Submission
1185	GATCGGTTTTTGCAGACATTTGGATTCTCACAAGAGATTGTGCTTACAAAGACAAAGTG	1244	Submitted (01-DEC-1998) Arabidopsis thaliana Genome Center,
330	LeuProLeuIleGlyAsnIleLysAlaProGlnValProIleThrCysIleMetGlyThr	349	Department of Biology, University of Pennsylvania, 38th Street and
1245	TTGCGCTTAACAGACAGACGATGATGCTCCGGAAGTCCGACGTCACTTCATATATGGGAGA	1304	Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
350	GlyValIleGlyThrLeuGluThrLeuPheTrpGlyLysGlyAspPheAspGluArgProGlu	369	6 (bases 1 to 99690)
1305	GGAGCTGATGATCCGGAGGTTTGTATGTATGAAAGAGATTCGATTAAGCAACAGAG	1364	Chen, K., Shin, P., Brooks, S., Buehler, E., Chao, Q.,
370	IleSerTrpGlyAspGlyAspGlyThrValAsnLeuValSerLeuLeuAlaLeuGlnSer	389	Johnson-Hopson, C., Khan, S., Kim, C., Altafi, H., Bel, B., Chin, C.,
1365	ATTAGTATGAGATGAGATGAGATGAGCGGTTAATTTGGCGAGCTTACAGAGCTTTG	1418	Chou, J., Choi, E., Conn, L., Conway, A., Gonzalez, A., Hansen, N.,
390	LeuTrpLysGluGluLysAsnGlnTrpLeuLysValValLysIleAspGlyValSerHis	409	Hwang, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, J.,
1419	-----AAAGTCGATAGCTTGACACCGCTAGAGATGTGAGATTGGAGTTTGGCAT	1463	Liu, S., Mukharzky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H.,
410	ThrSerIleLeuLysAspGluValAlaLeuAsnGluIleValGlyGluIleTrpSerIle	429	Schwartz, J., Southwick, A., Thaveri, A., Toriumi, M., Vaysberg, M.,
1464	ACATCTTATTACTTAAAGACGAGATCGCACTTAAGATATATGAACGAGATTTCATTAAT	1523	Yu, G., Davis, R., Federspiel, N., Theologis, A. and Ecker, J.
430	Asn 430		Direct Submission
1524	AAAT 1526		Submitted (15-AUG-2000) Arabidopsis thaliana Genome Center,
LOCUS	AC004557	99690 bp	Department of Biology, University of Pennsylvania, 38th and
DEFINITION	Genomic sequence for Arabidopsis thaliana BAC F17L21 from		Hamilton Walk, Philadelphia, PA 19104-6018, USA
ACCESSION	AC004557		On Aug 9, 2000, this sequence version replaced g1:3228389.
VERSION	AC004557.2	GI:9755398	
KEYWORDS	HTG.		1. 99690
ORGANISM	Arabidopsis thaliana.		/organism="Arabidopsis thaliana"
REFERENCE	Arabidopsis thaliana.		/db_xref="taxon:3702"
AUTHORS	Eukaryota: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta;		/chromosome="1"
	Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots;		/clone="F17L21"
	Rosidae: eurosids II: Brassicales; Brassicaceae; Arabidopsis.		complement(join(71..672,759..858,941..1801))
	1 (bases 1 to 99690)		/note="similar to transportin gl AAC25708.1; similar to
	Johnson-Hopson, C., Brooks, S., Buehler, E., Chao, Q., Khan, S., Kim, C.,		Est gb N96069.1"
	Shin, P., Altafi, H., Bel, O., Chin, C., Chou, J., Choi, E., Conn, L.,		/codon_start=1
	Conway, A., Gonzales, A., Hansen, N., Hwang, B., Koo, T., Lam, B.,		/evidence=not_experimental
			/product="F17L21.1"
			/protein_id="AAF99750.1"
			/db_xref="GI:9802548"
			/translation="MSVHGRRFPASPIISLSPSSSTSPSSQSPNPDPDLKORVIACLN
			KADROTIALASAEIDSIKRNLTTHDSFSEFLNCINTHTSSSKYSPKRCVALLSYLS
			YHDSLTPHLAKWSTVIRRLDPPDSVSACAVATVAMSADVTRQPAQVAKPLIEIT
			LIOEDSNLQIGALICLASVDAADPESEQRLSKLPKIGKLKSDGKRAAALLSAVA
			GSITTAGAGTKVLDMLVPLIEFLSEEDMARSKSAEALCKVATIEDLASQKTKQ
			TTALSREREDPKSVRETKRNRLALNLKMEVSTDDESLSPSRSDDGNGIGCFSSVTR
			STIDVGLKSKAPKRYPIKRRPSLPLPVNSVYATQKKNLPRKNGNMNTMIVEASSS
			DNKGPHFTPVKSSSETEKANSNGDPDIITKHTSKRSBDSVSFGGLRSGSRVAP
			SDGDSVKNCKDDEESKSDSELSLIRQALILENQQSLDLQVYTWSLISIVSR
			YVKGLEGYECCR"
			join(7691..7830,7965..8089,8764..8922,9121..9329)
			/note="hypothetical protein"
			/codon_start=1
			/evidence=not_experimental
			/product="F17L21.2"

/protein_id="AA99722.1"
 /db_xref="GI:9802520"
 /translation="MAGEGSKPATVDDAAYVIRTVSTFPHNDKDYDEMAIMKNK
 ARKIDNTCEIEYKELIKGRDLISGFNAPLPCLEIADMYNI EYORKSNMDDVO
 KAIRDSYEMIPLOODTORRRPNDRSRGQWOKPIHGFIKCDNRTVESLGNTRNEL
 HNLIIEIQHWRKFSNAIIKWSFRSTNKADRLAKELENNVTFSHSTG"
 join(10122, .10261, 10378, .10624, 10700, .10801, 10890, .10961)
 /note="hypothetical protein"
 /codon_start=1
 /evidence=not_experimental
 /product="F17L21.3"
 /protein_id="AA99723.1"
 /db_xref="GI:9802521"
 /translation="MYERARVQPEPTLSDAHSYITAAVEAFDEPTKEEFIKLNDIR
 DHGVDAAGIAKTELKIGHPRLRGISFEFPQVNDIHEAKRTIILKOKATIPPEA
 AYGAASYTKIKOIEPDWENFMNMLKTRRSIDLTVHVESFLKIMIMYDGKSEKVEY
 QEEVDLLIYHEDLIDFRLFMK
 join(12453, .12613, 12758, .12879, 13398, .13444, 13777, .13826,
 13874, .13939, 14016, .14285, 14382, .14499, 14614, .14703,
 14781, .14852)
 /note="hypothetical protein"
 /codon_start=1
 /evidence=not_experimental
 /product="F17L21.4"
 /protein_id="AA99724.1"
 /db_xref="GI:9802522"
 /translation="MVGSGSKHYEGESKPRATKDDAVYLRAVRAKPHNDKKYDDFV
 AVMTNKRARDCIKKEYEOLKGRDLISGFNAPLPCLEIKNYIRGAGDELYVITA
 KRKLARSDSKVFTFMLEAHRFRDEPAKROFISLDRARIDATFVGVGLV
 ELIKHDLILGFNALLPARFOIPIPIAGQVNVGRVPEPTI EATSYLNSKRAF
 HDERPAVEELKILNDIEARVDAASIASVEEDHQTLNGFSVFLSAEMKFIK
 LKGDGSHVAVSVQILRMVSEGNKSKSEAFQEVPLVQDHEDLVMELIKIIFSDP"
 join(15962, .16104, 16188, .16481, 16578, .16695, 16782, .16883,
 16966, .17034)
 /note="hypothetical protein"
 /codon_start=1
 /evidence=not_experimental
 /product="F17L21.5"
 /protein_id="AA99725.1"
 /db_xref="GI:9802523"
 /translation="MVGSKSPFNEDGNAYVDKVALQDTEPEKYQFVRIPIQY
 TANRFEETLSASLOELRDHANICGSNVLFPRLQTAIPPSASAFHINKVGRSV
 PPAVAPTMADATSYNAVKEAFHDEPAKMETKITLTDKARINASVIAEMELL
 KDHNLILGFCVLSPTBREITKAKRFLDGSQVVDVLIIRMSEGNKSKDEASQ
 EVRALIIGHEDLIMELSEIFSDP"
 join(17912, .17988, 18081, .18335, 18450, .18639, 18726, .18827,
 18928, .19014)
 /note="hypothetical protein"
 /codon_start=1
 /evidence=not_experimental
 /product="F17L21.6"
 /protein_id="AA99726.1"
 /db_xref="GI:9802524"
 /translation="MEEVHDEPEKFNHLLIRDFNHHRDARITACMGEIMRDL
 NLVREFDEPAESGLAOLQIATSVSPDSTIDAVSYIMVKEAFHDEPAKVEY
 FQLEFDIRLIDVAGITRREELIAKKNLIVLIMFLPEARITILKTEQAAYSI
 NKRKVAASRTGLKERQGDHRYVESFLEITLMYQEGKSVNDLTOEGHEDLVMPRA
 DVKRPDPGSSSARDS"
 join(19604, .19831, 20016, .24392)
 /note="similar to polypeptide db|BAA78425.1"
 /codon_start=1
 /evidence=not_experimental
 /product="F17L21.7"
 /protein_id="AA99727.1"
 /db_xref="GI:9802525"
 /translation="MEVFDKQTQCDPKPIPCSETEFGEDYKTLCLDGKPHVGPVVD
 KTRILADGGKRAVLAMPNIIHSLKIMICAESEFKSLINLAATLEFIIHYESSES
 MASEVITSDSTSLVNMNTKILSNIMLRROYOALINCYDILGYDGSITVPR
 ATTANGAVTVNPAFKHQRQDOLITISALGALISISQPIISRTTSAELITWLMQDY
 AKPSWSHIOOLROOIKOMKDKTSIDEFFQGLVLRDQALLKPMESAEYQMEVIEG
 LSDYKQVLDIOGREVPSLIEHLKHEVLAALASLSISAAASYRPPANK
 HNSNMYRGQNRNNRNGANSYQOQPRNDGSGYQKCOICGVFSARCSOLOMS
 GAVSTSPSQYPRATVWQPRANMAAMSVYRWLLDSGATHLLTDDNLNLALHQPYNG

EYVIAADGSTLPIITHGSSSTLSQSRSLALNLIYVNLKMLISVYKLCNANKVSE
 FPPAFQVNDLSTGARLQGRTKDELEYEMVPSNTPIELASPTPKTLPSMHSRLQH
 PSPPVKLSIVSOFSLPYVNSGOKHPCSHLKHSLPYSMTIISTPRLVYISVY
 WTSVTSVNDVFKYLLIYDITRYTMYLPIKQSOVRETVAKALVENFOKIRTL
 YSDNGEFLARQGLTHGISHLSLSTHEHNGIARKHRLHLEGLTLTQASTPT
 SYMTYAGFAIVYLINLRPSVNLNNEYSFKLFRSPVYLKRVFGCSFPMRLYTNH
 KLERSSQPCVFLQSAYLCLDRSGGVYTSRHQFVEDQPFESIDTHSVNSNS
 PEASPCQHPQSRIPLOSAPPLVQWSSPLPSOSHRPNAETSSSSSSNDVY
 VSKDNOVDNRNPFQTSSSSAOSNNSPSSITONPNPSPSPSPSPSPSPSPSP
 SSTSATSVNPPPPPTNNHPRKTRAKHITPRKLSILAKTVTRQIPIPTVNO
 ALDRKWRNAGEEINQIRNNTFELVPPPNQVISTKMIFTLKYPLNCTLDYKAR
 LVARGFROQYGLHYSEFSPVVSLLTRVLQDLAVSWSMILQDLVNNALQGLTYR
 VYVQPPGFIDPDRPHVCRKLKALYGLKQAPRAMYQELNPFQSLGSLADTDFE
 VYINDIOIVCYVVDIISRMKIDAKPVSTMAHPRKLSISGLALDERGRTVIGSL
 SOGILHMTKVVYVDLISRMKIDAKPVSTMAHPRKLSISGLALDERGRTVIGSL
 OYLAFTPRDIAVNRISQMHPRPTDIHQWAKRYLRLAGTATHGLILNSPESLH
 AFSDADWAGNDDEVSTNAYIVYGSPIAMSSKQGVAVRSSTEAYRAVANTSTI
 RWVCSLTLEGLTLPKMPVYICNVGATVYLSANVFHSRKHALLDHFTRDVSAGA
 LRVSHISTHQLDALTKRPLRQHFLOFSSKIGVSKLPSPS"
 25676, .26104
 /note="unknown protein: similar to ESTs db|AAV549501.1,
 db|AAV53434.1, db|AAV537554.1, and db|AAV521645.1"
 /codon_start=1
 /evidence=not_experimental
 /product="F17L21.8"
 /protein_id="AA99728.1"
 /db_xref="GI:9802526"
 /translation="MKRQDRQGSRYVYDLSVLNLRSPMIPISLPDHPPPARI
 RPPSSHRFSFSSSLHISPSGFSALLIISVALMKGSGVTTFIGLLMPWALALIL
 VFYVAGIVSAISWGRSLICVLTLPSPSPSSISGNVGEV"
 join(27209, .27349, 27618, .28079)
 /note="unknown protein: similar to ESTs gb|A1995662.1, and
 db|AAV549408.1"
 /codon_start=1
 /evidence=not_experimental
 /product="F17L21.9"
 /protein_id="AA99729.1"
 /db_xref="GI:9802527"
 /translation="MESSEDEVNLSRAIEKLHKKRREASGAFIEDADQFLSSL

Alignment Scores:

Pred. No.: 3e-103 Length: 99690
 Score: 1310.00 Matches: 264
 Percent Similarity: 62.57% Conservative: 62
 Best Local Similarity: 50.67% Mismatches: 92
 Query Match: 56.15% Indels: 104
 Gaps: 6

US-09-857-612a-14 (1-443) x AC004557 (1-99690)

QY 1 MetLysLysGluGlnGluGlyLeuLysIleGluValAlaThrLeuThValThrVal 20
 Db 88492 ATGAAAAA-----ATATCTCACATTAATTCGATGCAATACGATCTC 88536
 QY 21 ValValValMetLeuSerLeuLeuGlyThrCysGlyAlaSerAnLeuAspLeuLeu 40
 Db 88537 GTTGGTGGACGATGACCTCGATGCTGACCTGGTGGTACCACTGTTGATT 88596
 QY 41 LeuIleProGlyAsnGlyLysGlnLeuGluAlaAlaArgLeuThrAsnGlnTyrLysPro 60
 Db 88597 CTGGTTCAGGAACGGAGGATACAGTACGACGCTGGAGCAGAGATACAGCA 88656
 QY 61 SerThrPheIleCysGluSerTrp--TyrProLeuIleLysLysLysAsnGlyTyrPhe 79
 Db 88657 AGTAGTCTCTGCTAGCAGCTGCTTATTCGATTCAATAGAGAGAGTGGATGCTT 88716
 QY 80 ArgLeuThrPheAspSerSerValIleLeuAlaProPheThrGlnCysPheAlaGluArg 99
 Db 88717 AGGCTATGCTTCGATGACGACGATGATTATCTCCCTTACACAGAGTCTTACGATGCA 88776
 QY 100 MetThrLeuHisTyrGlnGluLeuAspAspTyrPheAsnThrProGlyValAlaThr 119
 Db 88777 ATGATGTTGTACTATGACCCCGATTGATGATTACCAAAATGCTCCGATGCTCAAAC 88836

QY	120	ArgValProHisPheGlySerThrHisSerLeuLeuTyLeuAsnProArgLeuLysHis	139
Db	88837	CGGGTTCTCATTTGGGTTCCGACCAAAATCATTCTATACCTCGACCTTGCTCCGGTT-	888935
QY	139	-----	139
Db	88896	AGTACTTCCAAAGATATATCATTTTGGGACATTTGCATTAATGACAAATATACATTAAT	889555
QY	139	-----	139
Db	88936	TTGGGGGATATTATTGTTATATCATATCCATTTTATATGCTAGTCGGTAATGTGATGTTAT	890155
QY	139	-----	139
Db	89016	GTTACTATATGTAATGTGAGTGTATGTGATTTTCCATTTTAAATGAAGCTGAAGATTGG	890755
QY	139	-----	139
Db	89076	TGCTTTAATTAATGTTGCTATGTCATGAGAAATTATAGGACACTATGTAATGTAGCTTAA	891355
QY	140	-----	151
Db	89136	TAATAAGGTTTGATTTGGCAGAGATCCACATCTTACATGGAACATTGGTGAACCTCTA	891935
QY	152	Gln---LysLeuGlyTyrAlaAspGlyLuthrLeuPheGlyAlaProTyrAspPheArg	170
Db	89196	GAGAAAAATGGGGGTATGTTTACCAACCAAAACATCTAGAGAGCTCCATATGATTTTCAGG	892555
QY	171	TyrGlyLeuAlaAlaGlyHisProSerGlnValGlySerLysPheLeuLysAspLeu	190
Db	89236	TACGGCTTGCGTCTTCGGGGCCACCCGTCGCTGTAGCTCACAAGTCTCTACAAACCTC	893155
QY	191	LysAsnLeuIleGlnLysAlaSerAsnSerAsnAngLysProValIleLeuLeuSer	210
Db	89316	AAACAAATGTTGGTGAATAAACTAGCAGCAAGCAAGCAAGAGCCAGTAGTACTCTCTCC	893755
QY	211	HisSerLeuGlyGlyLeuPheValLeuGlnLeuLeuAsnAtgAspProProSerTrpArg	230
Db	89376	CATACCTTAGGGAGGACTTTTGCTCTCTCATTTTCTCAACCGTACACCCCTTCATAGGGGC	894355
QY	231	LysLysPheIleLysHisPheIleAlaLeuSerAlaProTrpGlyAlaAlaLeuAspGlu	250
Db	89436	CGCAGTACATCAAAACACTTTGTTGTCACCTGGCGGCATGGGGGTGGGAGCATCTCTCAG	894955
QY	251	MetTyrTrpPheAlaSerGlyAsnThrLeuGlyValProLeuValAspProLeuLeuVal	270
Db	89496	ATGAAGCAATTTGCTTCTGGCAACACACTCGGTGCCCTTAGTAAACCTTCTCTGGTGC	895555
QY	271	ArgAspArgGlnArgSerSerGlySerAsnLeuTrpLeuLeuProAsnProLysIlePhe	290
Db	89556	AGACGGCATCCAGAGACCTCCGAGAGTAAACCAATGGCTACTTCCATTCACCAAAAGTGT	896155
QY	291	GlyProGln---LysProIleValIleThrProIleArgProTyrSerAlaHisAspMet	309
Db	89616	CACGACAGAACTAAACCCTTGTGCTACTACCTCCCGAGTTCATCTACACAGACTTACAGATG	896755
QY	310	ValAspPheLeuLysAspIleGlyPheProGlnGlyValIlyrProTyrGlyLuthrArgIle	329
Db	89676	GATCGGTTTTTTCGACAGATATGGATGTCACAAAGAGATGTGCTTTACAAAGCAAGAGTGG	897355
QY	330	LeuProLeuIleGlyAsnIleLysAlaProGlnValProIleThrCysIleMetGlyThr	349
Db	89736	TTGGCTTTAACAGAGAGACTGATGATCTCGGGAGTGGCCAGTCACTTGATATATGGGAGA	897955
QY	350	GlyValGlyThrLeuGlyLuthrLeuPheTyrGlyLysGlyAspPheAspGluArgProGlu	369
Db	89796	GGAATTTGATACACCGAGAGCTTTTGATGTATGTAAGAAAGAGATTCGATAAACCAACAGAG	898555
QY	370	IleSerTyrGlyAspGlyAspGlyThrValAsnLeuValSerLeuLeuAlaLeuGlnSer	389
Db	89856	ATTAAAGTTCAGATGAGATGGAGCGGTTTAATTTATGGCCAGCTTATGACAGCTTTG----	899095

Qy	390	LeuTprplysUGLUGLlyAsnGlnGlntryrlyysValaLysrllleAspGlyValserHis	409
Db	89910	-----AAAGTCGATAGCTTGGAACACCGTAGAGATGATGATTCGAT	89954
Qy	410	ThrsrllleLeuLysAspGluValaLeuAsnGlnIuleValGlyIuleThrsrllle	429
Db	89955	ACATCTATACCTTAAGACGAGATTCGACTTAAGACGATTATGACGATTCATTAATT	90014
Qy	430	Asn 430	
Db	90015	AAT 90017	
RESULT 8			
LOCUS	AC120983/c		
DEFINITION	Oryza sativa chromosome 3 clone OSUNB0011H13, *** SEQUENCING IN		
ACCESSION	AC120983		
VERSION	AC120983.1		
KEYWORDS	HTG: HTGS_PHASE2.		
SOURCE	Oryza sativa.		
ORGANISM	Oryza sativa		
REFERENCE	Eukaryota: Viridiplantae; Streptophyta: Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		
AUTHORS	1 (bases 1 to 134188) Buell,C.R., Yuan,Q., Ouyang,S., Liu,J., Tallon,L., Gansberger,K., Kim,M., Overton II,L., Bera,J., Tsalirli,T., Krol,M., Jarrahl,B., Jin,S., Koo,H., Zismann,V., Hsiao,J., Blunt,S., Vanaken,S., Uutterbach,T., Feldblyum,T., Yang,Q., Haas,B., Suh,B., Peterson,J., Quackenbush,J., White,O., Salzberg,S. and Fraser,C.		
TITLE	Oryza sativa ssp. japonica cv. Nipponbare OSUNB0011H13 BAC genomic sequence		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 134188)		
AUTHORS	Buell,R.		
JOURNAL	Direct Submission		
COMMENT	Submitted (14-MAY-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA * NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces * is believed to be correct as given, however the sizes * of the gaps between them are based on estimates that have * provided by the submittor. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved. 1 61218: contig of 61218 bp in length * 61219 61256: gap of unknown length * 61257 77652: contig of 16396 bp in length * 77653 77722: gap of unknown length * 77728 131160: contig of 53433 bp in length * 131161 131197: gap of unknown length * 131198 134188: contig of 2991 bp in length. location/Qualifiers 1.134188 /organism="Oryza sativa" /cultivar="Nipponbare" /db_xref="taxon:4530" /chromosome="3" /clone="OSUNB0011H13" /note="japonica cultivar-group" BASE COUNT 39144 a 28013 c 28230 g 38642 t 159 others ORIGIN		
Alignment Scores:			
Pred. No.:	3,55e-91	Length:	134188
Score:	1174.00	Matches:	234
Percent Similarity:	65.11%	Conservative:	72
Best Local Similarity:	49.79%	Mismatches:	121
Query Match:	50.32%	Indels:	44
GB:		Gaps:	5

/product="Putative ubiquitin / ribosomal protein CEP52"
 /protein_id="AA019122.1"
 /db_xref="GI:20330759"
 /translation="MOIPVKTLTGKTIILEVESSDTIDNVAKITQDEKGIPTPDQORLLI
 FAGQLEGRATLADNVIOKESTLHLVLRGCIIEPSLQALARKYNDOKMCRNEFKR
 SLRSLSPGPEAFSEEMQDKLMDGLMDLRLVDSRIFSCVARKLRPRAVNCRRKKCGHS
 NOLRRKKIKTK"
 complement(11559..12591)
 /gene="OJ1175C11.3"
 /note="Putative hemoglobin 4"
 complement(join(11559..11705,12033..12149,12260..12374,
 12467..12591))
 /gene="OJ1175C11.3"
 /codon_start=1
 /product="Putative hemoglobin 4"
 /protein_id="AA019123.1"
 /db_xref="GI:20330760"
 /translation="MAFASASNGAVRFTIEEDALVKSMAIMKDSANIGRFLKLFIE
 EVASARHLFSLNSDVPLEKNPKLKHAAVFMVTCFAAOLRKGRTVVRDTTIL
 RIGSHFHNKGVSDHFEVAREPALLETTIKDGIAPASWSPEMKNANGAEVYEHVIAIKEGS
 MKPVALL"
 complement(15258..16098)
 /gene="OJ1175C11.4"
 /note="Putative hemoglobin 3"
 complement(join(15258..15401,15490..15606,15700..15814,
 15965..16098))
 /gene="OJ1175C11.4"
 /codon_start=1
 /product="Putative hemoglobin 3"
 /protein_id="AA019124.1"
 /db_xref="GI:20330761"
 /translation="MAANGSNVSRGAVRFTEQEDALVKSMAIMKDSAHIGHREFE
 KTFEVSAPAROLFSLNSDVPLEKNPKLKHAAVFMVTCFAAOLRKGRTVVRDTTIL
 KIGSHFHNKGVSDAHFEVAREPALLETTIKEAVPASWSPAMKGANGAEVDHLVAIHK
 KQGMPPAA"
 complement(18978..19804)
 /gene="OJ1175C11.5"
 /note="Putative Non-symbiotic hemoglobin 1"
 complement(join(18978..19118,19234..19350,19446..19560,
 19677..19804))
 /gene="OJ1175C11.5"
 /codon_start=1
 /product="Putative Non-symbiotic hemoglobin 1"
 /protein_id="AA019125.1"
 /db_xref="GI:20330762"
 /translation="MAIVEDNAAVVSFESEOPALVTKSNAIIRKDSANILRFLFKI
 FEVASAOSMPSFLNSDVPLEKNPKLTHKMSVFMVTCFAAOLRKGRTVVDITDIL
 KRUGVTHLKYGGADHFEVAREPALLETTIKKEVPADWSPAMKSAMSEAYDHLVAIKQ
 EKRPAAE"
 complement(20266..20413)
 /note="Similar to Tourist_011 MITE element"
 complement(20216..23919)
 /gene="OJ1175C11.6"
 /note="Hypothetical protein"
 complement(join(20716..21204,21467..21553,21790..21889,
 21965..22371,22488..22925,23037..23153,23496..23579,
 23869..23919))
 /gene="OJ1175C11.6"
 /codon_start=1
 /product="Hypothetical protein"
 /protein_id="AA019126.1"
 /db_xref="GI:20330763"
 /translation="MDGRTDSSPRCNVOHQLNAEIKDNSTNRKLRPTPELPSLOLI
 EVOHLEKRLNDFANRRALERKALGKPPATIHSSNESCIPKTEBELIKETIANLELEVICIO
 LQHLHLAKRAFDQICSVSSCDMEINKOSARFSGLITGSSBLDSTYRKIQLDL
 SSGWMAKRSPTTLTJSTRTSHVNDKIGRSHSLGIRISARVSPSANNLARLAL
 KCHRLPLSPFEQKCGMPGIVSLADIIGTRIDAHVPTPKISEDMTCIASIYIRIRI
 RPNVVOHPPEFSPCSSFSASGLSSKTYGDIMSPRCRKEGYIEAMODDALTGESRYS
 FSOQDVIYEYALCKGAORADVDVMDHKKSYQLLESLDINGMKNEKTIAMVWV
 HNAAMMHLSYLISQGRVNPBELIYHILCCRVHSPQIOLRLILYRKMSKEEDIGQRRV
 VDRPELVHFAISGSSHDPAVRIRTPRRLIQOLEAARDEEVRANVGVGGRRKRGRRR
 VALLPLKLEIEPSRDAGADHLTRAVESCLEPELRPAADQAAARSRGGGGVEMRPHNH
 FAFRLRLRELIVGPAFPAHLSTST"

repeat_region
20811..20828
/note="Simple Sequence Repeat (CGC)n"
repeat_region
21305..21425
/note="Similar to SINEIOS Rice p-SINE1 retroposon"
gene
28171..30050
/gene="OJ1175c11.7"
/note="Hypothetical protein"
join(28171..28347,28968..29037,29138..29228,29490..2982961..30050)
/gene="OJ1175c11.7"
/codon_start=1
/product="Hypothetical protein"
/protein_id="AAI19127.1"
/db_xref="GI:20330764"
/translation="MAAANLAHAAPPPEPVLARAVAGAPERKRIEYVWERENTLSEFGR
AAVAIAAFKFNVOENDLTCRROFTGTRONAOARGNGFNSOGCDEVEPEO
ADVAIEFGICTMHLNMLRAISEKRSIDMETETKGMADKKAIADVDYEGNPBYN
TSSKKACVDBGKSTSDAETDDEGVCPKEKCAIAEKRSISDAETDDEGVPRNLHM
KDLDIMEPTTYQDHVKCTNRYSRYGGQTSTEFVEMDP"
complement(30482..33215)
/gene="OJ1175c11.8"
/note="Contains similarity to transposase"
complement(join(30482..31066,31149..31870,31963..32043,
32139..32213,32948..33215))
/gene="OJ1175c11.8"
/codon_start=1
/product="Putative transposase"
/protein_id="AAI19128.1"
/db_xref="GI:20330765"
/translation="MEEEYEIPDLNDPGVOEYLQDSGCIPLNLVPVAGCEBAFO
VEDELDPNOCFGABEDHEHPAMQAVELNSGRSAOEICHLNMPANVGDDIVFDGD
EYFLQDSNVADEDEDVVSYPYADVDEDLHTSSDEPTOODNIKRKYLEPAEKK
ATVALASTINGKLARDMTETIIAAMDYLRVYODITWTKAKCLAAGEVDFSKK
PGNGOKRSDIDLOVALRLNLSRCSIASLASLANASKSOVHRLYEGALRRHSNSTK
PYLEANKRKORLEPCVSMDLAATILNAEPRTIDMONIYIHDEKNFNATKKTYYLPN
EDEVRCKVSNKSRNEDAKFWFLSTLAKPRYDEGCNFEDGKIGIWPFVARKPEAOSRRR
PKGLTVKSIMNSRETISRARELTIKVIPALIASCWREDPVGTIMVODNALTHLPDE
AFALAIVQSGLDIRINMQPNSPMDNLDGFASIQSKTYLKNCISMDDLINVEE
YKENSINQSVRFLLNQACEFLEVKAADGNQKIPHNKDRLERLMDLPSLNCIAL
YNKMQLLN"

repeat_region
complement(30482..33215)
/note="Contains similarity to transposase"
complement(34294..37827)
/gene="OJ1175c11.9"
/note="Contains similarity to RINC zinc finger protein"
complement(join(34294..34436,34580..34704,34769..35317,
36029..36509,37120..37396,37528..37827))

CDS

Alignment Scores:

Pred. No.:	2.24e-49	Length:	123620
Score:	696.50	Matches:	190
Percent Similarity:	35.51%	Conservative:	71
Best Local Similarity:	25.85%	Mismatches:	154
Query Match:	29.85%	Indels:	320
DB:	8	Gaps:	11

US-09-857-612A-14 (1-443) x AC103891 (1-123620)

OY 18 ValThrValValValMetLeuSerLeu-----LeuCysThyCsglyAla 33
||| ||||||| ||||| ||||||| |||
Db 66677 GTCGCCGTTGAGCCGTCCTGCTCGCGTCGCGTCGCGTCCGCCGCCGCGCGCGC 66736
||| ||||||| ||||| ||||||| |||
OY 34 SeranLeuAspProLeuIleLeuIleProGlyasnGlyGlnLeuGluAlaArg 53
||| ||||||| ||||| ||||||| ||| ||||| |||||
Db 66737 AGCGATGTCACACCCTGCTGCTGCGCGGGAGTACAGCGACTCTACCGCGG 66796
||| ||||||| ||||| ||||||| ||| ||||| |||||
OY 54 LeuThrAsnGlnTylysProSerThrPhelIecysGluSerTrpTyProLeuIleLys 73
||| ||||||| ||||| ||||| ||||| ||||| |||||
Db 66737 CTGACTCGCGCGCTACGACCGCGCGCGCTCGCTCGCGCGG -----CGACAG 66844
||| ||||||| ||||| ||||| ||||| ||||| |||||
OY 74 LysLysasngLyTrpPharGlyLeuTrp---PheaspSerSerValIleLeuAlaProPhe 92
||| ||||||| ||||| ||||| ||||| ||||| |||||
Db 66845 GGGAAGCATMGATGCTTGACGCTGCGCGATGACGCGCGCGAGCGAACCGCGCAG 66904

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
US-09-857-612a-14 (1-443) x AP005640 (1-139534)
Qy 13 ValaIathrleuThrValThrValValValMetleuSerleu----- 28
Db 128592 ATGGCACTATTCACCTGGTGGTAGTACAGCTGCTCTGCTGCTTCCACTTCCCTC 128651
Qy 29 -----CysThrCysGlyAlaSerAsnLeuasp 37
Db 128652 CGAGAGCACTCTGGTCAGCGCAACCGCGGAACGACGTCGATCGCGGCGAGCTCCAC 128711
Qy 38 ProleuIleuIleuIleuProGlyAsnGlyIleAsnGlnLeuGluAlaArgLeuThrAsnGln 57
Db 128712 CCGATCTCTGCTGTCGCCCGGGGTCGATGCGACGACCTGAGCGCGGCTCAGCAGAGCG 128771
Qy 58 TyrlsProserThrPheIleCysGluSerTrpTrpProleuIleuIleuIleuIleuAsnGly 77
Db 128772 TACCGCGCGTCGGCGGCGAGGTGGCGCG-----ATGAAGGGGAAGGCG 128816
Qy 78 TrpPheArgLeuTrpPheAspSerValIleuAla---ProPheThrGlnCysPhe 96
Db 128817 TGGTTCGGCGTGTGGAGAAACGCTCCGACCTTCGACGCAACCGCTCAACAGAGTGTTC 128876
Qy 97 AlaGluIleuMetThrLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 116
Db 128877 CTGAGAGAGATGACCTCATCTACGACCCCTCCGAAACGACTACCGGAACCTTCCCGCGC 128936
Qy 117 ValGluThrArgValProHisIleuGlySerThrAsnSerLeuLeuTyIleuAsnProArg 136
Db 128937 GTCCAGACCGCGGTGCCCTACTGCTGCTC----- 128966
Qy 137 LeuIleuHisIleuThrGlyTy----- 143

Db 128967 -----GTCAAAGCTACCATCAGAAATGCGCTTTTGACAAAGTAATCTCCGCCCA 129017
Qy 143 ----- 143
Db 129018 AATTAAACGCTTACATATGCTGCTCACTGCTACTGCCGTGATGATGATGATGATGCT 129077
Qy 143 ----- 143
Db 129078 TCTCTTAGTTTTAACTGCTGCTGATCAAGTATCATGCTCGATCGATGATGATGCTGCT 129137
Qy 144 -----MetAlaProleuValAspSerLeuGlnIleuGlyTyAla 157
Db 129138 GCTTCACAGCCATGCTGCTTACACCGCTTATACAGCCCTGGAAGATGGATACCGT 129197
Qy 158 AspGlyIleuThrLeuPheGlyAlaProTyIleuAspPheArgTyIleuAlaIleuGly 177
Db 129198 GACGGGCAACATGACAGCGAGCTCTTACATTCCTCATGTTCTCCAGTCCCTGCG 129257
Qy 178 HisProserGlnValGlySerIleuPheLeuIleuIleuIleuIleuIleuIleuIleu 197
Db 129258 CAGGATCCCAAGTCTACTCGCGCTACTACGAGAGTTCATGAGACTGCTGGAGGCCCA 129317
Qy 198 SerAsnSerAsnAsnGlyIleuProValIleuLeuSerHisSerLeuGlyIleuPhe 217
Db 129318 AGCAAGAGCGATTCGCAAGAGAGATGATCATTTCTCGGCGACAGCCAGGAGATGCGTC 129377
Qy 218 ValLeuGlnLeuLeuAsnArgAsnProProSerTrpArgIleuIleuIleuIleuIleu 236
Db 129378 GCCCTTGGATGTCGTGAGAACACCCCTGGGTGGGAGAAAGATATCATAGCAACCTC 129437
Qy 237 PheIleAlaLeuSerAlaProTrpGlyValAlaIleuAspGluMetTyThrPheAlaSer 256
Db 129438 TTCTGCTGACGCTACCTCTCCCGCGGCTGCTCGACCCGGTGGAGAACCTTCCAC 129497
Qy 257 Gly-----AsnThrLeuGlyValProleuValAspProleuValArgAspGluGln 274
Db 129498 GGGCCGCAACACCTGCTTCTAGCTGCACAGCGAGCGAGCGTCTGCTCGGCCAATGTGG 129557
Qy 275 ArgSerSerGluSerAsnLeuTrpLeuLeuProAsnProIleuIleuPheGlyProGlnIleu 294
Db 129558 AGGACCTTGTGAGACGACGATGCCAACCTCCGTCGCCGCGGTGTTGGG---CGCAG 129614
Qy 295 ProIleValIleuThrProIleArgProTySerAlaHisAspMetValAspPheLeuIleu 314
Db 129615 CCGATCTGTGTCACCGACGCGAGGACTACTTCGCGTACGATATGGAAGATCTCTTCC 129674
Qy 315 AspIleGlyPheProGluGlyValTyIlePro----- 324
Db 129675 GCGCTGCGTTCGTCGACGGGATCGAGCCTTACAGTAGGATGCAACGGGATTTTTT 129734
Qy 324 ----- 324
Db 129735 TCCGCTTATCCCACTTCTAGTTATTTTTTCTTAAATTTTTGTACTACATGTGGGAA 129794
Qy 325 -----Tyr 325
Db 129795 TAACTACCAAGAGATTTTATATGCGGGTAGTGGGACTACACCATTTGCATCTTTC 129854
Qy 326 GluThrArgIleuLeuProleuIleuGlyAsnIleuIleuIleuIleuIleuIleuIleu 345
Db 129855 AGGACAGCAATGTTGTAGGATGTAACCTCGAGGGCGCCATGTGCGGTGACGAT 129914
Qy 346 IleMetGlyThrGlyValGlyThrLeuGluThrLeuPheTyIleuGlyIleuAspPhe 365
Db 129915 ATCATGTGAGAGGGGAGAGAGCGACCGCAATGTGCTATGTGGAGCAACTTTCAC 129974
Qy 366 GluArgProGluIleuSerTyIleuGlyAspGlyTrpValAsnLeuValSerLeuLeu 385
Db 129975 AAGGCTCCGGAAGAGTATAGTATGATGAGATGGGATAGTTAACTGTGTTACCAATGTTA 130034
Qy 386 AlaLeuGlnSerLeuTrpIleuGluIleuIleuIleuIleuIleuIleuIleuIleuIleu 403

```

db 130035 GCGTTCGACGAGCAATGCCTGGCACCAGACAGAGGGCCCATTTCAAGTCATCAATCAAG 130094
Oy      404   lIeaSpGIvAlSeRlHtSrSerlLeuYAspGIvAlAlaLeuAmsnLIuIleVAl 423
          :::::  |||||  |||||  |||||  |||||  :::::
Db 130095 GTCGAGAACCGCTCCCATGTGGAAATTCTTGATGAATGCGCTTAAGAAGGTCATG 130154
Oy      424   GIyGCIUIEtHrSerlIeAsn 430
          |||||  ::::|

Db 130155 CAAGAGATTCTTGAAAGTCMAT 130175

RESULT 12
AP004073 LOCUS AP004073 147548 bp DNA linear PLN 13-MAR-2002
DEFINITION Oryza sativa [japonica cultivar-group] genomic DNA, chromosome 1,
PAC clone:p0614D08.
ACCESSION AP004073
VERSION AP004073.2 GI:19386859
KEYWORDS
SOURCE
ORGANISM
Oryza sativa [japonica cultivar-group] (cultivar:Nipponbare) DNA,
clone:p0614D08.
Oryza sativa [japonica cultivar-group] Embryophyta; Tracheophyta;
Eukaryota; Viridiplantae; Streptophyta; Liliopsida; Poales; Poaceae;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eubratiodideae; Oryzaceae; Oryza.
REFERENCE
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:p0614D08
JOURNAL Published Only In Database (2001)
REFERENCE 2 (bases 1 to 147548)
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
JOURNAL Direct Submission
Submitted (15-AUG-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tasakientas@affrc.go.jp, URL:http://irgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Mar 12, 2002 this sequence version replaced gi:15208441.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI Nonredundant Protein database, nr
(tftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RCR. Protein homologies of the coding regions were searched against
NCBI Nonredundant protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DBJ accession no. and RCP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from SP6 to T7 of the PAC clone.
This sequence of p0614D08 clone has an overlap with P0660F12
(DBBJ:AP003687) clone at the position 1 to 135,297 of 5' end. The
sequence of this clone starts at the position 11,907 of P0660F12.
Detailed information on overlap and assembly quality together with
annotation of this entry is available at
http://irgp.dna.affrc.go.jp/GenomeSeq.html.
Location/Organism
1..147548
/organism="Oryza sativa [japonica cultivar-group]"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="P0614D08"
/complement(3037..3535)
/gene="P0614D08.1"
/complement(3037..3535)
/gene="P0614D08.1"
/note="probably inactive due to frameshift(s) in CDS

```

	pseudogene"
gene	/pseudo complement(join(7123..7412,7452..7583,8141..8285, 8890..9097,10827..11083)) /gene="P0614D08.2" complement(join(7123..7412,7452..7583,8141..8285, 8890..9097,10827..11083)) /gene="P0614D08.2" /note="hypothetical protein" /codon_start=1 /protein_id="BAB86237.1" /db_xref="GI:19386860" /translation="MPAGAWLPPTGDMKTTNDGALSGGGSSGGSHTITRAGAARRR GSGDAEESASILMTHAHRSRGMRTPAVLAMYWRCAARSROPDSADSGDLRDGR SARVYIERGRRLIGVGCRKRREERERSPATIHVEEQAPRPDSAAANRELGDAA TILEVEDPERLTMTMSADEGEVARTMKRTTGIVTHRDGTAKOVARREKKDDPRS RYRPGEDSREROANOYVTSDGEGYGCGGCPMARAGVSYTYTKATROCEPDREKAC ROEBTEDEPANRLSGSGGGGKRRLLGIRSAATELGLTRPSGPRRRNSDGAAGETTD VASKITRST"
CDS	complement(join(11424..11791,11799..12120)) /gene="P0614D08.3" complement(join(11424..11791,11799..12120)) /gene="P0614D08.3" /note="hypothetical protein" /codon_start=1 /protein_id="BAB86238.1" /db_xref="GI:19386861" /translation="MDWVKRHMGSHGKGARGEEAPGSDGSSORMAGGERROACGAR HKREGLKAVPTSGDEGEAARMRRCCGGMGTATATRRPAVAASGDSLGLKAGAR RIEGFDNRKRAERAGACVASCGRKKKEGQAGGEGRIDGPVTVIDARKREQS RRRVARRPCGGRRRRFLAAPAVAKKCEELTERRGRTPPRRGAEMARFGKDS TAASAALKQG"
gene	complement(join(12756..12787,13386..13509,14014..14210, 16440..14763,15312..15440,19176..19286)) /gene="P0614D08.4" complement(join(12756..12787,13386..13509,14014..14210, 16440..14763,15312..15440,19176..19286)) /gene="P0614D08.4" /note="hypothetical protein" /codon_start=1 /protein_id="BAB86239.1" /db_xref="GI:19386862" /translation="MRNSSRSEGDQSIILEMEEITYVEDSELKKKKNGKEIQNTPMAN BDSGLASVTLTWPFSPSENKDEDEVYVIKOGGINAMIKRIHHYTKYSQGSTDI TFEGGTQCSCKPKIESSCSCSGPEMTTDTDBEANNNSSGTNGLPNTGTGNRR ELGRNLNHFTSTSPSELAEIYTIARWMSPSGSNISGDLGSLRQISAKNTGLPY ANERGPGTGSSTRIFEKSA"
misc_feature	20178..22247 /gene="P0614D08.5" 20178..22247 /gene="P0614D08.5" /note="probably inactive due to frameshift(s) in CDS pseudogene similar to Oryza sativa chromosome 10, OSUBA0046L02.1"
gene	/pseudo join(25104..25136,25949..26227,26431..26850) /gene="P0614D08.6" join(25104..25136,25949..26227,26431..26850) /gene="P0614D08.6" /note="contains EST AU102097(C10257) unknown protein" /codon_start=1 /protein_id="BAB86240.1" /db_xref="GI:19386863" /translation="NSTSLPAEEKKAARKKLKGMDFPIFRNFYWRVSVLGQLSV SNGADESMCLRHPQRNSVVNESSLNHNISFLGRIOQLSTGHESRRYSTAFRNML TKMWAAYNHRSAKYAAHVACGLHKVYANHVTTKLEFKYMLSTLVNDEQAQAVCCI RIEPTREWNICSPACTCTRFNSLSINTEGYKCAKCHFSIASISTYIRIMILECYW ATMOPEKLAKKHONCHVSIRKLL"
CDS	join(30179..30310,30384..30591,31666..31796) /gene="P0614D08.7" join(30179..30310,30384..30591,31666..31796) /gene="P0614D08.7" /gene="P0614D08.7"

Db 135577 GTTGATCGGTGATGACCCGATCAGGTTACTTCAAGTCATTTTGTATCCCAACAGCAG 135636
Qy 408 rHlStrSerIleuLeuLysAspDilaAlaLeuAsnGluIleValaIGluIleThrse 428
Db 135637 GCACAAGGTCATCTCGCATGATTAACGCTCAAGCGTGTGTGACAGCAAAATTTCTTG 135696
Qy 428 rIleAsnSer 431
Db 135697 AGCAATTC 135706
RESULT 13
E26773
LOCUS E26773 1236 bp DNA linear PAT 18-JUN-2001
DEFINITION Novel protein and DNA thereof.
ACCESSION E26773
VERSION E26773.1 GI:13020711
KEYWORDS JP 199269199-A/8.
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1236)
AUTHORS Kao, T.
TITLE Novel protein and DNA thereof
JOURNAL Patent: JP 199269199-A 8 05-OCT-1999;
TAKEDA CHEM IND LTD
OS Unidentified
PN JP 199269199-A/8
PD 05-OCT-1999
PF 10-APR-1998 JP 1998099664
PR
PI KAO TANIYAMA
PC C07K14/47, A61K31/70, A61K38/45, A61K38/45, A61K48/00,
PC A61K48/00,
PC C07K16/40, C07K19/00, C12N1/19, C12N1/21, C12N5/10, C12N9/10, PC
C12N9/09, C12N5/09,
PC C12P21/08, C12Q1/48, G01N33/566//C12N1/19, C12R1:865), (C12N1/21,
PC C12R1:19),
PC (C12N1/21, C12R1:125), (C12N5/10, C12R1:91), (C12N9/10, C12R1:19),
PC A61K37/52,
PC A61K37/52, A61K37/52, C12N5/00, C12N15/00, (C12N5/00, C12R1:91) CC
Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..1236
FT /organism='Unidentified'.
FEATURES
source 1..1236
location/Qualifiers
BASE COUNT 269 a 342 c 348 g 277 t
ORIGIN
Alignment Scores:
Pred. No.: 2.15e-44 Length: 1236
Score: 612.00 Matches: 139
Percent Similarity: 54.55% Conservative: 83
Best Local Similarity: 34.15% Mismatches: 161
Query Match: 26.23% Indels: 24
DB: 6 Gaps: 10
US-09-857-612a-14 (1-443) x E26773 (1-1236)
Qy 16 LeuThrValThrValValValMetIleuSerLeuLeuCysThrCysGluYlaSerAsn 35
Db 49 CTCCTGTACCTCTGTTTCTACTAATGATGCTGCAGACCTGACGCTCCCGCCACACT 108
Qy 36 LeuAspProLeuIleLeuIleProGluAsnGlyGluAsnGluLeuGluAlaArgLeuThr 55
Db 109 CACCCCGCGTGTGCTGCTGCTGCTGATTTGGGTAACCACTGGGAACCAAGCTGAT 168
Qy 56 AsnGluTyrLysProSerThr-----PheIleCysGluSerTrpTyrProLeuIleLys 73

Db 169 -----AAGCCAAAGGTTGTACACTACTTGTCTC-----AAG 201
Qy 74 LysLysAsnGlyTrpPheArgLeuTrpPheAspSerSerValIleuAlaProPheThr 93
Db 202 AAGACGAGACAGCTACTTACACACTGTGGCTGAATCTGGAACTGCTGTGCTGTATCAT 261
Qy 94 GluCysPheAlaGluArgMetThrIleuHisTyrHisGluGluLeuAspTyrPheAsn 113
Db 262 GACTGCTGATTCACAAATATCAGCGTGTGTACACAGCAACATCTCGGCCACCTGTT 321
Qy 114 ThrProGluValGluThrArgValProHisPheGlySerThrAsnSerLeuLeuTyrLeu 133
Db 322 CCCGATGTTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 381
Qy 134 AsnProArgLeuIleuHisIleThrGlyTyrMetAlaProLeuValAspSerLeuGluLys 153
Db 382 GACCCGAGCAAGAGAAATGTTGGTTCCTTTTCTACACATATGTGTGAGAGCTTGTGGG 441
Qy 154 LeuGlyTyrAlaAspGlyLeuThrLeuPheGlyAlaProTyrAspPheArgTyrGlyLeu 173
Db 442 TGGGCTACACACAGCGGTGAAGACGTTGAGGTGCTCCTATGATGTGCGCGAGCC--- 498
Qy 174 AlaAlaGluGlyHisProSerGluValGlySerLysPheLeuLysAspLeuLysAsnLeu 193
Db 499 -----CCAAATGAAGGCGCTTACTTCTTGGC---CTGCGAGAGATG 540
Qy 194 IleGluGluAlaSerAsnSerAsnGlyLysProValIleLeuLeuSerHisSerLeu 213
Db 541 ATCGAGAGCATGTACCAAGATGATAGGGGCG---CCGCTGTCTGTGCTGCCACAGCAG 597
Qy 214 GlyGlyLeuPheValLeuGluLeuAsnArgAsnProProSerTrpArgLysLysPhe 233
Db 598 GGCACGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 657
Qy 234 IleLysHisPheIleAlaLeuSerIleArgProTyrGlyAlaIleAspLysMetTyrThr 253
Db 658 ATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 717
Qy 254 PheAlaSerGlyAsnThrLeuGluValProLeuValAspProLeuLeuAlaArgAspGlu 273
Db 718 CTGGCTCAGAGACACAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 777
Qy 274 GluArgSerSerGluSerAsnLeuTrpLeuLeuProAsnProLysIlePheGlyProGlu 293
Db 778 CAGCGATCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837
Qy 294 LysProIleValIleThrProIleArgProTyrSerAlaHisAspMetValAspPheLeu 313
Db 838 AAGGTAATTTGTATACACACACCACTAAGTACAGCGCTCCGGGACTATACCCGTTCTTC 897
Qy 314 LysAspIleGlyPheProGluGlyValTyrProTyrGluThrArgIleLeuProLeuIle 333
Db 898 CGGACATGCGTTTTCGAAGATGCGTGTTCATCGCGCAGAC---ACAGAACGGCTGTT 954
Qy 334 GlyAsnIleLysAlaProGluValProIleThrCysIleMetGlyThrGlyValGlyThr 353
Db 955 GAAGCATGACCCACCCGCGGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
Qy 354 LeuGluThrLeuPheThrGlyLysGlyAspPheAspGluArg---ProGluIleSerTyr 372
Db 1015 CCAAACTCTTCTACTAC-----GAGACCTTCTGATGCGGACCCCAAAATGTGCTTC 1068
Qy 373 GlyAspGlyAspGlyThrValAsnLeuValSerLeuLeuAlaLeuGluSerLeuTyrLys 392
Db 1069 GCGATGTGTGAGCGGAGCTGTAAGCTGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1125
Qy 393 GluGluLysAsnGluTyrLeuLysValValLysIleAspGlyLysSerHisThrSerIle 412
Db 1126 ACCCCCAAGACACAGATATCATTCGAGAGCTGCGGGAAGGAGACACATTCAGATG 1185
Qy 413 LeuLysAspGluValAlaLeu 419
Db 1186 CTAGCCAAATGCCACCACTTG 1206

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT
MGC:25713	BC019373	BC019373.1	house mouse.	Mus musculus	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus	Strausberg, R.	Submitted (19-DEC-2001)	National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	NIH-MGC Project URL: http://mgc.ncl.nih.gov	Contact: MGC help desk

```
|||||
Db 1157 GGCATGTGACGGCAGGTGAACCTGGAGAGCTCTCCAGATGCAAGCC---TGCAG 1213
QY 393 GluLysAsnGlnTyrLeuValValLysIleAspGlyValSerHisThrSerIle 412
Db 1214 AGCCGCCAAGAGCAGAGATATCATGTGAGAGCTGCCGGGAGCGAGACATTGAGATG 1273
QY 413 LeuLysAspGluValAlaLeu 419
Db 1274 CTAGCCAATGCCACCACTTG 1294
```

Search completed: February 15, 2003, 14:56:17
Job time : 2688 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2003, 13:14:57 : Search time 215 seconds

(without alignments)
4640.164 Million cell updates/sec

Title: US-09-857-612a-14

Perfect score: 2333
Sequence: 1 MKKEGEGKIEVATLTVV.....GETTSINSHAEISLNFSG 443

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framer_p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09857612/runtat_11022003_111817_29814/app_query.fasta_1.583
-DB=N.Geneseq_101002 -QPMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTMT=ptc -NORM=ext -HEAPSIZE=500 -MTNLEN=0 -MAXLEN=2000000000
-USBR=US09857612.ecgn_1.1.125_&runtat_11022003_111817_29814 -KCPu=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2333	100.0	1332	21 AAA49206	Soybean putative 1
2	2022	86.7	1217	21 AAA49202	Soybean putative 1
3	1380.5	59.2	1299	22 AAS01081	Arabidopsis thaliana
4	1380.5	59.2	1495	21 AAC47273	Arabidopsis thaliana
5	1380.5	59.2	1498	21 AAC39845	Arabidopsis thaliana
6	1310	56.2	1562	21 AAC64437	Arabidopsis thaliana
7	1310	56.2	1562	21 AAC64437	Arabidopsis thaliana
8	913.5	39.2	1500	21 AAA49204	Corn putative leci
9	739.5	31.7	1660	21 AAA49205	Corn putative leci
10	717	30.7	1440	21 AAA49203	Corn putative leci
11	612	26.2	1236	19 AAV68568	Murine kidney leci
12	610	26.1	1236	19 AAV68566	Human heart leci
13	610	26.1	2680	20 AA234023	Human PRO540 nucle
14	610	26.1	2680	21 AAC78492	Human PRO540 (UNO3
15	610	26.1	2680	21 AAA88516	Human PRO540 CDNA
16	610	26.1	2680	22 AAC91556	CDNA encoding huma
17	610	26.1	2680	24 ABK65967	Human cDNA sequenc
18	610	26.1	2681	22 AAH14618	Human full-length
19	610	26.1	2720	22 AAK94279	Human heart leci
20	608	26.1	1137	19 AAV68561	Human heart leci
21	608	26.1	1146	19 AAV68564	Human heart leci
22	606	26.0	1137	19 AAV68563	Murine leci
23	597	25.6	2687	21 AAA26392	Human secreted pro
24	596	25.5	1332	19 AAV68567	Human kidney leci
25	594	25.5	1233	19 AAV68562	Human kidney leci
26	594	25.5	1242	19 AAV68565	Human kidney leci
27	560	24.0	1824	23 ABL21223	Drosophila melanog
28	554	23.7	1135	22 AAS01095	Corn sterol acyltr
29	543	23.3	1358	8 AAN70191	CDNA clone p12 enc
30	541	23.2	1323	24 ABK87900	Human encoding huma
31	541	23.2	1744	20 AA232180	Human leci
32	541	23.2	1744	22 AAD02648	Human leci
33	541	23.2	1744	24 ABN95747	Gene #2245 used to
34	541	23.2	1744	24 ABK35510	Human endometrial
35	541	23.2	4308	22 AAD02649	Plasmid for
36	493	21.1	563	21 AAA69662	Pinus radiata leci
37	490.5	21.0	4024	23 ABL21222	Drosophila melanog
38	458.5	19.7	921	21 AAA49201	Corn putative leci
39	428	18.3	264	22 AAS01087	Soybean sterol acy
40	376	16.1	353	21 AAA69561	Pinus radiata leci
41	372.5	16.0	542	21 AAA49200	Corn putative leci
42	344	14.7	273	22 AAS01088	Soybean sterol acy
43	281.5	12.1	5700	24 ABK87899	Human leci
44	281.5	12.1	6901	18 AAT78853	Human leci
45	248	10.6	778	22 AAH06918	Human cDNA clone (

ALIGNMENTS

RESULT 1
ID AAA49206 standard: CDNA: 1332 BP.

AAA49206;

12-DEC-2000 (first entry)

Soybean putative leci

Soybean: leci

phosphatidylcholine-sterol O'-acetyltransferase; heat shock; cold shock; ss.

Glycine max.

Key Location/Qualifiers
CDS 1..1332
FT /*tag= a

FT		/product= "lecithin:cholesterol acyltransferase"
FT		/EC_number= "2.3.1.43"
XX		
PN	WO200032791-A2.	
XX		
PD	08-JUN-2000.	
XX		
PF	02-DEC-1999:	99WO-US26586.
XX		
PR	03-DEC-1998:	98US-0110782.
PA	(DUPO) DU PONT DE NEMOURS & CO E I.	
PI	Cahoon RE, Kinney AJ, Sakai H, Shen JB, Butler KH, Saylor JJ;	
DR	WPI: 2000-412337/35.	
XX	P-PsDB: AAB01211.	
PT	Polyelectrode encoding plant lecithin:cholesterol acyltransferase	
XX	enzyme useful for producing transgenic plants and for producing	
PS	antibodies specific to which is useful for screening cDNA expression	
XX	libraries -	
PS	Claim 2: Page 46; 49pp; English.	
XX		
CC	The present sequence is a putative coding sequence for a soybean	
CC	lecithin:cholesterol acyltransferase (also known as	
CC	phosphatidylcholine-sterol O-acetyltransferase). This enzyme is found	
CC	associated with high-density lipoproteins and Apolipoprotein-AI and -D.	
CC	The gene and protein can be used to produce transgenic plants which have	
CC	increased lipid metabolism and membrane fluidity, and therefore increased	
CC	resistance to heat and/or cold shock, to alter the content of phytoestrob	
CC	or lecithin in grains and to identify potential herbicides.	
XX		
SQ	Sequence 1332 BP; 385 A; 282 C; 289 G; 376 T; 0 other;	
	Alignment Scores:	
	Pred. No.: 1,036-240 Length: 1332	
	Score: 2333.00 Matches: 443	
	Percent Similarity: 100.00% Conservative: 0	
	Best Local Similarity: 100.00% Mismatches: 0	
	Query Match: 100.00% Indels: 0	
	DB: 21 Gaps: 0	
US-09-857-612A-14 (1-443) x AAA49206 (1-1332)		
OY	1 MELLVSLVSGUUGLUGLUGLULGULYLYLSIIIEGLVALValThrLeuThrValThrVal	20
Db	1 ATGAAGACAGCAACAAGAAGGGCTCACAAGATTGAGGTGCACACTCACAGTTACAGTA	60
OY	21 VALValValMetLeuSerLeuLeuCyThrnCysGlyAlaSerAsnLeuAspProLeuIle	40
Db	61 GTTGTGTGATGCTGTGCATTCGTCATGCATCATGTGGGGCAAGCAACTCGACCTTGATT	120
OY	41 LeuIleProGlyAsnGlyLyasnGlnLeuGluAlaArgLeuThrAsnGlnTyrrLySpro	60
Db	121 CTAAATACCAAGGTAAAGGAGAACCAACATAAGCAAGTAGTGACCAATCGTACAAAGCCC	180
OY	61 SerTrpPheIleCysGlnSerTrpTyrrProLeuIleIleLySlysAsnGlyTrpPheArg	80
Db	181 TCTACTTTTCATCTGGCAATCATGTGATCCCTCTTCATCAAGAAAAAGATGGATGCTTCAGA	240
OY	81 LeuTrpPheAspSerSerValIleLeuAlaProPheThrGlnCysPheAlaGluArgMet	100
Db	241 CTTTGCTTTGATTCACAGTGTCTACTTGTCTCTTTTACTACATATGCTTTGCCGAAGCATG	300
OY	101 ThrLeuHisTyrrHisGlnLeuLeuAspAspTyrrPheAsnThrProGlyValGluThrArg	120
Db	301 ACCCTTCATTTACCACCAACAACACTCGATGATTACTTCAACACCTCGGGGCTTGAGACCCG	360
OY	121 ValProHisPheGlySerThrAsnSerLeuLeuTyrrLeuAsnProArgLeuLySHisIle	140
Db	361 GTCCCTGCACTTTGGTGGCCCAACATCTCTTCTATCTCAATCTCTGTCGAAGCATATC	420

QY	141	ThrGlyTyrMetAlaProLeuValAspSerLeuGlnIleuSglYtyrAlaAspGlyGlu	160
Db	421	ACCGGATACATGGACCCCTGGTAGATTCAATACAAAGCTGGCTACGCTGATGGTAG	480
QY	161	ThrLeuPheGlyAlaProTyrAspPheArgGlyrcGlyLeuAlaIleGlnGlnHisProSer	180
Db	481	ACTCGTGTGGAGCCCTTATGACTTATGATATGCTGCTAGCTGTAAGGTCAACCTTCA	540
QY	181	GlnValGlySerLysPheLeuLysAspLeuLysAsnLeuIleGlnGlnAlaSerAsnSer	200
Db	541	CAAGTGGGTCCAGATGCTCTCAAGATCTAAAGAAATTTGATPAGAGAGAACACCAATTC	600
QY	201	AsnAsnGlyLysProValIleLeuLeuSerHisSerLeuGlyLeuPheValLeuGln	220
Db	601	AATAATGGGAAGCCAGTATCTCTCTCCACAGTTTAGAGGCCATATTTGCTCTACAA	660
QY	221	LeuLeuAsnArgAsnProProSerTyrArgLysLysPheIleLysHisPheIleAlaLeu	240
Db	661	CTACTAATATAGAAACCCCCCTCTTGGGCAAAAAATTCATCAACCTTCATGTGCTT	720
QY	241	SerAlaProTyrGlyGlyAlaIleAspGlnuMetLyrThrPheAlaSerGlyAsnThrLeu	260
Db	721	TCAGCTCCATGGGGGGTGGTATGAGAGAAATGACACCTTGCATGTGGCAACACTTGG	780
QY	261	GlyValProLeuValAspProLeuLeuValArgAspGlnuArgSerSerGlySerAsn	280
Db	781	GGAGTGGCCCTAGTGGACCCCTTTTGTAGTGGGGATATACAAAGAGCTCGAGAGTAA	840
QY	281	LeuThrPheLeuProAsnProLysIlePheGlyProGlnLysProIleValIleThrPro	300
Db	841	CTTTGGCTTTTGCTTACACCCAAAAATTTTGGTCTTAAAAACCATAGTGAATCTCA	900
QY	301	IleArgProTyrSerAlaHisAspMetValAspPheLeuLysAspIleGlyPheProGln	320
Db	901	ATTAGCCCTTATGAGCTCATGACATGATGATTTTCTAAAGACATGGTTCCTGAA	960
QY	321	GlyValLyrProTyrGlyThrArgIleLeuProLeuIleGlyAsnIleLysAlaProGln	340
Db	961	GGGGTTATCCCTTATGAAACCGAATTCCTACCTTGATAGGAAACATAAAGCACACAA	1020
QY	341	ValProIleThrCysIleMetGlyThrGlyValGlyThrLeuGlnThrLeuPheTyrGly	360
Db	1021	GTGGCTAATAACTGTGATTTATGGAGACGGAGTGGGAACCTTGGAAACATTTGTTATGG	1080
QY	361	LysGlyAspPheAspGluArgProGlnIleSerTyrGlyAspGlyAspGlyThrValAsn	380
Db	1081	AAAGGTGATTTTGTATGAACGGCCCGAATATCATATGGGATGCTGTGAAACGGGTAA	1140
QY	381	LeuValSerLeuLeuAlaLeuGlnSerLeuTyrPheGlnGlnLysAsnGlnTyrLeuLys	400
Db	1141	TTTGGAGACTGTGGCGCTTCAATGCATATGGAAAGAGAGAAATCATACCTTTAA	1200
QY	401	ValValLysIleAspGlyValSerHisThrSerIleLeuLysAspGluValAlaLeuAsn	420
Db	1201	GTGGTTAAGATAGATGGGGGTGTCTCATACTTCAATTAAGATGAGTGGACATAAT	1260
QY	421	GluIleValGlyGluIleThrSerIleAsnSerHisAlaGluLeuGlyLeuSerHisLeu	440
Db	1261	GAATAGTAGGTGAGATTACTTCAATTAATTCATATGCTGAGCTCGTTTAAGTAATTTG	1320
QY	441	PheSerGly 443	
Db	1321	TTTTTCGGGG 1329	
RESULT 2			
AAA49202			
ID	AAA49202 standard; cDNA; 1217 bp.		
XX	AAA49202:		
XX			
DT	12-DEC-2000 (first entry)		

Oy	398	TyrLeuLysValIValLysIleAsnGlyValSerHisThrSerIleLeuLysAspGluVal	417
Dd	1165	AGCTTGAACACCACCGATGAGATTGCATTCATCATCTAATTAAAGCAGATC	1225
Oy	418	AlaLeuAsnGluIleValGlyIleThrSerIleAsn	430
Dd	1225	GCACTTAAGAGATTATGACACGACATTTCATTTATTAAT	1263
RESULT 4			
AAC47273	ID	AAC47273 standard; DNA: 1495 BP.	
XX	AC	AAC47273:	
XX	DT	18-OCT-2000 (first entry)	
XX	DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 53208.	
XX	KW	Hybridisation assay; genetic mapping; gene expression control;	
XX	KM	protein identification; signal transduction pathway;	
XX	OS	metabolic pathway; promoter; termination sequence; ss.	
XX	PN	Arabidopsis thaliana.	
XX	PD	EPI033405-AZ.	
XX	PD	06-SEP-2000.	
PF	PE	25-FEB-2000; 2000EP-0301439.	
PR	PR	25-FEB-1999; 99US-0121825.	
PR	PR	05-MAR-1999; 99US-0123180.	
PR	PR	09-MAR-1999; 99US-0123548.	
PR	PR	23-MAR-1999; 99US-0125788.	
PR	PR	25-MAR-1999; 99US-0126264.	
PR	PR	29-MAR-1999; 99US-0126785.	
PR	PR	01-APR-1999; 99US-0127462.	
PR	PR	06-APR-1999; 99US-0128234.	
PR	PR	08-APR-1999; 99US-0128714.	
PR	PR	16-APR-1999; 99US-0129845.	
PR	PR	19-APR-1999; 99US-0130077.	
PR	PR	21-APR-1999; 99US-0130449.	
PR	PR	23-APR-1999; 99US-0130510.	
PR	PR	23-APR-1999; 99US-0130891.	
PR	PR	28-APR-1999; 99US-0131449.	
PR	PR	30-APR-1999; 99US-0132048.	
PR	PR	30-APR-1999; 99US-0132407.	
PR	PR	04-MAY-1999; 99US-0132484.	
PR	PR	05-MAY-1999; 99US-0132485.	
PR	PR	06-MAY-1999; 99US-0132486.	
PR	PR	06-MAY-1999; 99US-0132487.	
PR	PR	07-MAY-1999; 99US-0132863.	
PR	PR	11-MAY-1999; 99US-0134218.	
PR	PR	14-MAY-1999; 99US-0134219.	
PR	PR	14-MAY-1999; 99US-0134221.	
PR	PR	14-MAY-1999; 99US-0134370.	
PR	PR	18-MAY-1999; 99US-0134768.	
PR	PR	19-MAY-1999; 99US-0134941.	
PR	PR	20-MAY-1999; 99US-0135124.	
PR	PR	21-MAY-1999; 99US-0135353.	
PR	PR	24-MAY-1999; 99US-0135629.	
PR	PR	25-MAY-1999; 99US-0136021.	
PR	PR	27-MAY-1999; 99US-0136782.	
PR	PR	28-MAY-1999; 99US-0137222.	
PR	PR	01-JUN-1999; 99US-0137528.	
PR	PR	03-JUN-1999; 99US-0137502.	
PR	PR	04-JUN-1999; 99US-0137724.	
PR	PR	07-JUN-1999; 99US-0138094.	
PR	PR	08-JUN-1999; 99US-0138094.	
PR	PR	10-JUN-1999; 99US-0138540.	
PR	PR	10-JUN-1999; 99US-0138847.	

PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148365.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 30-AUG-1999; 99US-0151080.
 PR 31-AUG-1999; 99US-0151303.
 PR 01-SEP-1999; 99US-0151438.
 PR 07-SEP-1999; 99US-0151930.
 PR 10-SEP-1999; 99US-0152363.
 PR 13-SEP-1999; 99US-0153070.
 PR 15-SEP-1999; 99US-0153758.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 06-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 08-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0158293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 22-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.: 3,07e-138
 Score: 1380.50
 Percent Similarity: 75.29%
 Best Local Similarity: 60.97%
 Query Match: 59.17%
 DB: 21

Length: 1495
 Matches: 264
 Conservative: 62
 Mismatches: 92
 Indels: 15
 Gaps: 5

US-09-857-612a-14 (1-443) x AAC47273 (1-1495)
 QY 1 MetLysLysGluGlnGluGluLysLLeGluValAlaThrLeuThrVal 20
 DB 82 ATCAAAAA-----ATATCTTCATATTATTCGGTAGCTACGATAC 126
 QY 21 ValValValMetLeuSerLeuLeuGlySerGlyAlaSerAsnLeuAspLeuIle 40
 DB 127 GTTGTCGTACGATGACCTCGATGTGTCAAGCTGGTGGTAGCAAGCTGACCTTGATT 186
 QY 41 LeuIleProGlyAsnGlyLysGlnGlnLeuGluAlaArgLeuThrAsnLntyrLysPro 60
 DB 187 CTGGTTCCAGAAACGAGAGTACACGCTAGAGTAGCGCTGCACAGCAAAATACAGCCA 246
 QY 61 SerThrPheIleCysGluSerTrp--TyrProLeuIleLysLysLysAsnGlyTrpPhe 79
 DB 247 AGTAGTGTCTGGTGTAGCAGCTGGTTATATCCGATTCAATACAGAGCTGGTAGTGT 306
 QY 80 ArgLeuTrpPheAspSerSerValIleLeuAlaProPheThrGlnCysPheAlaGlyArg 99
 DB 307 AGCGTATGGTGTGATGACAGAGTGTATGTGCTCCCTCACAGGTGCTTCAGCGATGCA 366
 QY 100 MetThrLeuHisTyrHisGlnGluLeuAspAspTyrPheAsnThrProGlyValGluThr 119
 DB 367 ATGATGTGTACTATGACCCCTGATTGGATGTATACAAAATGCTTCGTCTCAAAACC 426
 QY 120 ArgValProHisPheGlySerThrAsnSerLeuLeuThrLeuAsnProArgLeuLysHis 139
 DB 427 CGGTTCTCATTTTCGGTTCGACCAATACCTTCTTACTCTGACCCCTGCTCCGACAT 486
 QY 140 IleThrGlyTyrMetAlaProLeuValAspSerLeuGln--LysLeuGlyTyrAlaAsp 158
 DB 487 GCACATCTTACATGGAACATTTGGTGAAGCTCTAGAAAATAAGCGCGATGTATAC 546
 QY 159 GlyGluThrLeuPheGlyValProTyrAspPheAlaGlyTyrGlyLeuAlaIleGluGlyHis 178
 DB 547 GACCAAAACCATCTAGACCTCATATGATTTAGAGTAGCGCTGCTCTCGGCGCCAC 606
 QY 179 ProSerGlnValGlySerLysPheLeuLysAspLeuLysAsnLeuIleGluGluAlaSer 198
 DB 607 CCGTCCCGTGTACCCCTACAGATTCTCACAAGACCTCAACATTTGGTGAATAAACTAC 666
 QY 199 AsnSerAsnAsnGlyLysProValIleLeuLeuSerHisSerLeuGlyLysLeuPheVal 218
 DB 667 ACGAGAAACGAAGAAAGCAGTAGTACTCTCCCATAGGCTAGAGAGCTTTTCCGC 726
 QY 219 LeuGlnLeuLeuAsnArgAsnProProSerTrpArgLysLysPheIleLysHisPheIle 238
 DB 727 CTCATTTCTCAACCGTACCAACCCCTTCATGCGCCGCAAGTACATCAACACTTTGTT 786
 QY 239 AlaLeuSerAlaProTrpGlyGlyAlaIleAspGluMetTyrThrPheAlaSerGlyAsn 258
 DB 787 GCACCTGCTGCCCATGGGCTGGAGAGATCTCTCAGATTAACATTTCTTCTGGCAAC 846
 QY 259 ThrLeuGlyValProLeuValAspProLeuLeuValArgAspGluGlnArgSerSerGlu 278
 DB 847 ACACCTGGTGGCTTTAGTAAACCTTGTGTCAGACGCGCATCAGACGACCTCCGAG 906
 QY 279 SerAsnLeuTrpLeuLeuProAsnProLysIlePheGlyProGln---LysProIleVal 297
 DB 907 ACTAACCAATGGCTACTTCATCCAAAGCTTTCACGACAGCACTAACCGCTTGTC 966
 QY 298 IleThrProIleArgProThrSerLeuHisAspMetValAspPheLeuLysAsnGly 317
 DB 967 GTAACCTCCAGGTTAACTACAGCTTACGAGATGATGATGCTTTTTCGACACATGGA 1026
 QY 318 PheProGluGlyValTyrProTyrGluThrArgIleLeuProLeuIleGlyAsnLys 337
 DB 1027 TTCTCAAGAGAGGTGTGCTTACAAAGCAAGAGTGTTCCTTAAACAGAGAGCTGATG 1086
 QY 338 AlaProGlnValProIleThrCysIleMetGlyThrGlyValGlyThrLeuGluThrLeu 357

Db 1087 ACTCCGGAGTCCAGTCACTTCATATATGCGAGAGAGATTGATACACCGGAGCTTTG 1146
Qy 358 PheTyrglyIysGlyAspPheAspGluarProGluIleSerTyrglyAspGlyAspGly 377
Db 1147 ATGTATGGAAGAGAGATTTCGATACACACAGATTAACTATGCGAGATGCG 1206
Qy 378 ThrValAsnLeuValSerLeuLeuAlaLeuGlnSerLeuTrpLysGluGluLysAsnGln 397
Db 1207 ACGGTATATTGGCGAGCTTACAGCTTTG-----AAAGTCGAT 1245
Qy 398 TyrLeuLysValValLysIleAspGlyValSerHisThrSerIleLeuLysAspGluVal 417
Db 1246 AGCTTGACACCGCTTACGATGAGATTTCGATCATCTATATCTAAAGACGAGATC 1305
Qy 418 AlaLeuAsnGluIleValGlyGluIleThrSerIleAsn 430
Db 1306 GCACCTTAAGAGATTATGACGACATTTCAATTATTAT 1344
RESULT 5
AAC39845
ID AAC39845 standard; DNA; 1498 BP.
XX
AC AAC39845;
XX
DT 17-Oct-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 26113.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.


```
Db      967 GTACTCCCGCAGTTACTACACAGCTTACGAGATCGCTTTTTCAGACATTGCA 1026
      967      ::::|||||      ||:::||||:|||||      |||      |||||
Qy      318 PheProGluGlyValIYrProTyrGluThrArgIleLeuProIleGlyAsnIleYs 337
      1027 TTCCTCAAGAGGATGTGGCTTACAGCAAGAGTGTCCCTTACAGAGAGGACCTGATG 1086
      1027      |||      ::::|||||      |||||      |||||      |||
Qy      338 AlaProGluValIProIleThrCysIleMetGlyThrGlyValGlyThrLeuGluThrLeu 357
      1087 ACTCCGGGAGTCCAGTCACCTTCATATATGCGAGAGAGTGTGATACCGGAGCTTTTG 1146
      1087      |||      |||||      |||||      |||||      |||
Qy      358 PheTyrGlyGlyGlyAspPheAspGluArgProGluIleSerTyrGlyAspGlyAspGly 377
      1147 ATGATGGAAGAGGAGATTCGATACCAACAGAGATTAGTATGAGATGAGATCGG 1206
      1147      |||||      |||||      |||||      |||||      |||||
Qy      378 ThrValAsnLeuValSerLeuLeuAlaLeuGlnSerLeuTyrPlyGluGluLeuAsnGln 397
      1207 AGGCTTAAATTTTGGCGACTTACGACCTTGG-----AAAGTCGAT 1245
      1207      |||||      |||||      |||||      |||||
Qy      398 TyrLeuLysValValIleValIleAspGlyValSerHisThrSerIleLeuLysAspGluVal 417
      1246 AGCTTGAACACCGCTAGAGATGATGCACTTCGCATACATCTTATCTTAAAGACAGATC 1305
      1246      |||      ||:::|||||      |||||      |||||      |||||
Qy      418 AlaLeuAsnGluIleValIleGlyIleThrSerIleAsn 430
      1306 GCACCTTAAGAGATTATGAGACAGATTTCATATTAT 1344
      1306      |||||      |||||      |||||      |||||
```

RESULT 6

AAC64437 standard; DNA: 1562 BP.

AAC64437;

12-FEB-2001 (first entry)

Arabidopsis thaliana PDAT genomic DNA SEQ ID NO:10.

PDAT: phospholipid:diacylglycerol acyltransferase: triacylglycerol;

TAG; EST: expressed sequence tag; fatty acid; oil content; ds.

Arabidopsis thaliana.

WO200060095-A2.

12-OCT-2000.

28-MAR-2000: 2000WO-EP02701.

01-APR-1999: 99EP-0106656.

10-JUN-1999: 99EP-0111321.

07-FEB-2000: 2000US-0180687.

(BADT) BASF PLANT SCI GMBH.

Dahlgvist A, Stahl U, Lenman M, Banas A, Ronne H, Styhme S;

WPI: 2000-665012/64.

Phospholipid:diacylglycerol acyltransferase enzymes in the biosynthetic

pathway for triacylglycerol production and DNAs encoding them, useful

for producing triacylglycerol, or for transforming any cell or organism

to increase oil content -

Claim 6: Page 59; 97pp: English.

The present invention describes an enzyme for catalysing (in an

acyl-CoA-independent reaction) the transfer of fatty acids from

phospholipids to diacylglycerol in the biosynthetic pathway for the

production of triacylglycerol (TAG). The enzyme is designated as

phospholipid:diacylglycerol acyltransferase (PDAT). The enzyme and

the nucleotides encoding them are useful for producing TAG and/or

TAG with uncommon fatty acids. The enzyme and the nucleotide are

also useful for transforming any cell or organism in order to be

CC expressed in this cell or organism and result in an altered, preferably
CC increased oil content of this cell or organism. The present sequence
XX represents the Arabidopsis thaliana PDAT genomic DNA.
SQ Sequence 1562 BP; 441 A; 312 C; 367 G; 442 T; 0 other;

Alignment Scores:

Score:	1,24e-130	Length:	1562
Best Local Similarity:	62.57%	Matches:	264
Query Match:	50.67%	Mismatches:	92
	56.15%	Indels:	104
		Gaps:	6

US-09-857-612a-14 (1-443) x AAC64437 (1-1562)

```
Qy      1 MetLysLysGluGlnGluGluGlyLeuLysIleGluValAlaThrLeuThrValThrVal 20
      1      |||||      |||      |||      |||      |||
Db      1 ATGAAAAA-----ATATCTTCACATTAATTCGATGATACGATACCTC 45
      1      |||||      |||||      |||||      |||||
Qy      21 ValValValMetLeuSerLeuLeuCysThrCysGlyAlaSerAsnLeuAspProIle 40
      46 GTTGGTGACGAGTACCTCGATGTGTACAGCTGGGTGACACAGTGTACCTTGTGAT 105
      46      |||||      |||||      |||||      |||||
Qy      41 LeuIleProGlyAsnGlyGlyAsnGlnLeuGluAlaArgLeuThrAsnGlnTyrLysPro 60
      106 CTGGTTCAGAAACGAGGAGTACACAGTACAGTACGCTGGAGCAGAGATACAGCA 165
      106      |||||      |||||      |||||      |||||
Qy      61 SerThrPheIleCysGluSerTrp--TyrProLeuIleLysLysLysAsnGlyTyrPhe 79
      166 AGTAGGTCTGTGTGACGACGCTGCTTATATCCGATTAAGAGAGAGTGGATGGTT 225
      166      |||||      |||||      |||||      |||||
Qy      80 ArgLeuTrpPheAspSerSerValIleLeuAlaProPheThrGlnCysPheAlaGluArg 99
      226 AGGCTATGCTCGATGACGAGTGTATTGTCTCCCTTCACAGGCTTACAGCATGCA 285
      226      |||||      |||||      |||||      |||||
Qy      100 MetThrLeuHisTyrHisGlnGluLeuAspAspTyrPheAsnThrProGlyValGluThr 119
      286 ATGATGTTGTACTATGACCGCTGATTTGGATGATTACCAAAATGCGTCGGTCCAAAC 345
      286      |||||      |||||      |||||      |||||
Qy      120 ArgValProHisPheGlySerThrAsnSerLeuLeuTyrLeuAsnProArgLeuHis 139
      346 CGGTTCCCTCATTCGTTGCGACCAATCACTCTATACCTGACCCGTCGCGTT- 404
      346      |||||      |||||      |||||      |||||
Qy      139 ----- 139
      405 AGTACTTCCAAAGATATATCATTTTGGACATTTGCAATATGACAAATAGACATAAT 464
      405      |||||      |||||      |||||      |||||
Qy      139 ----- 139
      465 TTGGGGGATTATGTTATATCATATATCCATTATATGCTAGCGTAATGAGGTAT 524
      465      |||||      |||||      |||||      |||||
Qy      139 ----- 139
      525 GTTACTATAGTTAATGTGAGTGTATGTGATTTTCCATTTAATGAAGTAGAAGTTG 584
      525      |||||      |||||      |||||      |||||
Qy      139 ----- 139
      585 TCGTTTAATATGTTGCTATGTCATGAGAAATTAAGACACTATGTAATGTACCTTA 644
      585      |||||      |||||      |||||      |||||
Qy      140 ----- 151
      645 TAATAAGTTGATTTGACAGAGATCCACATCTTACATGGAAACATTTGTGAAGCTCTA 704
      645      |||||      |||||      |||||      |||||
Qy      152 Gln---LysLeuGlyTyrAlaAspGlyGluThrLeuPheGlyAlaProTyrAspPheArg 170
      705 GAGAAATAATCGCGGTATGTTAAGACCAACCACTACGAGAGCTCATATGATTTACAG 764
      705      |||||      |||||      |||||      |||||
Qy      171 TyrGlyLeuAlaIleGluGlyHisProSerGlnValIleGlySerLysPheLeuLysPhe 190
      765 TACGGCTGGGTGCTTGGGCGACCCGTCGCTGAGGCTACACATCTTACAGACCTCT 824
      765      |||||      |||||      |||||      |||||
Qy      191 LysAsnLeuIleGluGluAlaSerAsnSerAsnGlyLysProValIleLeuLeuSer 210
```

```
Db 825 AACAATTTGGTGAAGAACTACACGAGAACGAGAAAGCAAGTATCTCTCC 884
Oy 211 HAsSerLeuGlyLeuPheValLeuGlnLeuLeuAsnArgAnPProPSeRTpArG 230
Db 885 CAtAGCGCTAGAGACTTTTCCTCTCCATTTCTCAACCGTACCAACCTTATGCGC 944
Oy 231 LysLysPheLLeuLysHsPheLLeuLeuSerAlaProTrpLysLysAlaLeuGlu 250
Db 945 CCGAAATACATCAAAACACTTTGTCACCTCGCTCCATGGGGTGGAGACTCTCAG 1004
Oy 251 MeTTyTThPheLAsSerGlyAnThrLeuGlyValProLeuValAsPProLeuLeuVal 270
Db 1005 ATGAAGACATTTGCTTGGCAACACACCTGGCTCTTAACTTAACCTTGTGCTG 1064
Oy 271 ArgAspGlnLArgSerSerGlnSerAnLeuTrpLeuProAnProLysIlePhe 290
Db 1065 ACAGCGCATCAGAGACGACCTCCGAGATAACCAATGCTACTTCCATCTACCAAGTGT 1124
Oy 291 GlyProGln--LysProIleValLLeuThrProIleArgProTrpSerAlaHsAspMet 309
Db 1125 CACGACAGAACTAAACCGCTGTGCTGTAACCTCCAGGTTAACTACACAGCTTACGAGATG 1184
Oy 310 ValAsPheLeuLysAspLLeuGlyPheProGlnGlyValTyPProTrpGlnThrArgIle 329
Db 1185 GATCGGTTTTTGCACACATTTGATTCTCACAAGAGATTGCTTACAAAGACAGATG 1244
Oy 330 LeuProLeuLLeuLysAlaProGlnValProIleThrCysLLeuMetGlyThr 349
Db 1245 TTGCTTTACAGAGAGAGCTGATGATCCGGAGGTCCCATCTTGCATATATGGGAGA 1304
Oy 350 GlyValGlyThrLeuGlnThrLeuPheTyGlyLysGlyAspPheAspLuarProGln 369
Db 1305 GGAGTTGATACACCGGAGTTTGTATGTATGAAAGAGATTCGATTAAGCAACAGAG 1364
Oy 370 IleSerTyGlyAspGlyLysPglyThrValAsnLeuValSerLeuLeuAlaLeuGlnSer 389
Db 1365 ATTAAGTATGAGATGAGATGGAGCGTTAAATTGGCGACGTTACACCTTTG----- 1418
Oy 390 LeuTrpLysGlnGluLysAsnGlnTyLeuLysValIleLysLLeuAspGlyValSerHis 409
Db 1419 -----AAACTCCATAGCTTGAACACCGCTAGAGATTGATGAGTTTGCAT 1463
Oy 410 ThrSerIleLeuLysAspLysValAlaLeuAsnGlnIleValGlyLLeuThrSerIle 429
Db 1464 ACATCTATACTTAAGACGAGATCGACCTTAAGAGATTATGAGACGATTTCATTATT 1523
Oy 430 Asp 430
Db 1524 AAT 1526

RESULT 7
AAC64447
ID AAC64447 standard: DNA: 1562 BP.
AC AAC64447;
AC AAC64447;
DE 12-FEB-2001 (first entry)
XX Arabidopsis thaliana PDAT genomic DNA SEQ ID NO:9b.
XX Arabidopsis thaliana PDAT genomic DNA SEQ ID NO:9b.
XX PDAT: phospholipid:diacylglycerol acyltransferase; triacylglycerol;
KM TAG: EST; expressed sequence tag; fatty acid; oil content; ds.
XX Arabidopsis thaliana.
XX Arabidopsis thaliana.
XX WO200060095-A2.
PD 12-OCT-2000.
PF 28-MAR-2000; 2000WO-EP02701.
XX 01-APR-1999; 99EP-0106656.
```

```
PR 10-JUN-1999; 99EP-0111321.
PR 07-FEB-2000; 2000US-0180687.
XX (BADI ) BASE PLANT SCI GMBH.
PI Dahlqvist A, Stahl U, Lemman M, Banas A, Ronne H, Stryme S;
XX WPI: 2000-665012/64.
XX Phospholipid:diacylglycerol acyltransferase enzymes in the biosynthetic
PT pathway for triacylglycerol production and DNAs encoding them, useful
PT for producing triacylglycerol, or for transforming any cell or organism
PT to increase oil content -
XX
XX Claim 6: Page 93-94; 97pp; English.
XX
CC The present invention describes an enzyme for catalysing (in an
CC acyl-CoA-independent reaction) the transfer of fatty acids from the
CC phospholipids to diacylglycerol in the biosynthetic pathway for the
CC production of triacylglycerol (TAG). The enzyme is designated as
CC phospholipid:diacylglycerol acyltransferase (PDAT). The enzyme and
CC the nucleotides encoding them are useful for producing TAG and/or
CC TAG with uncommon fatty acids. The enzyme and the nucleotide are
CC also useful for transforming any cell or organism in order to be
CC expressed in this cell or organism and result in an altered, preferably
CC increased oil content of this cell or organism. The present sequence
CC represents the Arabidopsis thaliana PDAT genomic DNA.
XX
SQ Sequence 1562 BP; 441 A; 312 C; 367 G; 442 T; 0 other:

Alignment Scores:
Pred. No.: 1,24e-130 Length: 1562
Score: 1310.00 Matches: 264
Percent Similarity: 62.57% Conservative: 62
Best Local Similarity: 50.67% Mismatches: 92
Query Match: 56.15% Indels: 104
DB: 21 Gaps: 6

US-09-857-612A-14 (1-443) x AAC64447 (1-1562)
Oy 1 MetLysLysGlnGlnGlnLLeuLysLLeuGlnValAlaThrLeuThrVal 20
Db 1 ATGAAAAA-----ATATCTCAATTTCCGTAGCTATACCATCTC 45
Oy 21 ValValAlaMetLeuSerLeuLeuCysThrCysGlyAlaSerAsnLeuAspProLeuIle 40
Db 46 GTTGTGCTAGCATGACCTCGATGTGTCTAAGCTGTGGAGACAGACGTGACCTTTGATT 105
Oy 41 LeuLLeuProGlyAsnGlyLysAsnGlnLeuGlnAlaArgLeuThrAsnGlnTyLysPro 60
Db 106 CTGGTTCCAGAAACGAGACGTAACCAAGTACGCTGACAGCAAGAAATACAGCCA 165
Oy 61 SerThrPheLLeuCysGlnSerTrp---TyPProLLeuLLeuLysLysAsnGlyTrpPhe 79
Db 166 ACTAGTCTGCTGATGACGCTGCTGTTATCCGATTTCATAAAGAGAGTGGATGCTTT 225
Oy 80 ArgLeuTrpPheAspSerSerValLLeuAlaProPheThrGlnCysPheLLeuArg 99
Db 226 AGCGTATGCTTGCATGACGACACTGTTATGTCTCCCTTACCAAGTGTCTACGATCGA 285
Oy 100 MeTTThLeuHsITyThsLsGlnLLeuLysAspTyPTrpAsnThrProGlyValGluThr 119
Db 286 ATGATGTTGTACTATACCCCTGATTTGGATGTATACCAAAATGCTCTGCTCCAAAC 345
Oy 120 ArgValProHsPheGlySerThrAsnSerLeuLeuTyTrpLeuAsnProArgLeuHsHis 139
Db 346 CGGGTTCCTCATTTGCGTTGCGCAAAATCATTTATACCTGACCCCTGCTCGGTT- 404
Oy 139 ----- 139
Db 405 ACTACTTCCAAAGATATATCATTTTGGGACATTTTCATATACAAATAAGACATAAT 464
Oy 139 ----- 139
```


Db 465 TTGGGGATATATGTTATATCATATCCATTATATGCTAGTCGGTAATGAGTGTAT 524
 QY 139 ----- 139
 Db 525 GTTAGTATAGTAAATGATGATGTTATGATTTCCATTAAATGAAGTAGAAAGTTG 584
 QY 139 ----- 139
 Db 585 TCGTTATATATGTTGCTATCATGAGATTATAGAGACATATGTAATAGTACTTAA 644
 QY 140 ----- 140
 Db 645 TAATAGGTTGATTTGACAGATGCCATCTTACATGAAACATTTGGTGAACCTCTA 704
 QY 152 Gln---LysLeuGlyTyrAlaAspGlyLutThrLeuPheGlyAlaProTyrAspPheArg 170
 Db 705 GAGAAAAATCGGGGTATGTTAACGACCAACCATCTAGAGAGCTCCATATGATTTTCAGG 764
 QY 171 TyrGlyLeuAlaAlaGlyLysHisProSerGlnValGlySerLysPheLeuLysAspLeu 190
 Db 765 TACGGCCCTGGTCTGCTGGGGCCACCCGCTGAGCTTACACACTTCTTACAAAGACTC 824
 QY 191 LysAsnLeuIleGlyLysAlaSerAsnSerAsnAsnGlyLysProValIleLeuLeuSer 210
 Db 825 AAACAATGCTGTGAAAAAACTAGACGAGAGAGAAAGCAAGCCAGTATCTCTCC 884
 QY 211 HisSerLeuGlyGlyLeuPheValLeuGlnLeuLeuAsnAlaGlnProSerTyrArg 230
 Db 885 CATAGCCCTAGAGAGACTTTGCTCCATTTCTCAACCGCTACACCCCTTCAGGGCGC 944
 QY 231 LysLysPheIleLysHisPheIleAlaLeuSerAlaProTyrGlyAlaIleAspGlu 250
 Db 945 CGCAAGTACATCAACACTTTGTCACACTCCCTGGCCATGGGGGAGACATCTCTAG 1004
 QY 251 MetTyrThrPheAlaSerGlyAsnThrLeuGlyValProLeuValAspProLeuVal 270
 Db 1005 ATGAAGACATTTGCTTGGCAACACACTCGGTGCTTCTTAAAGTTCCTTGGTGTG 1064
 QY 271 ArgAspGluGlnArgSerSerGlnSerAsnLeuThrLeuProAsnProLysIlePhe 290
 Db 1065 AGAGCGCATCAGAGACTCCGAGAACCAATGGCTTCCATCCACCAAACTGTTT 1124
 QY 291 GlyProGln---LysProIleValIleThrProIleArgProTyrSerAlaHisAspMet 309
 Db 1125 CACGACAGAACTAAACCCCTTGTCTGTAACCTCCACAGTTACTACACAGCTTACAGATG 1184
 QY 310 ValAspPheLeuLysAspIleGlyPheProGluGlyValTyrProTyrGlyThrArgIle 329
 Db 1185 GATGGTATTTTGGACAGATGATTCACACAGAGAGTTGGCTTACACAGCAAGACTG 1244
 QY 330 LeuProLeuIleGlyAsnIleLysAlaProGlnValProIleThrCysIleMetGlyThr 349
 Db 1245 TTGCTTTAAACAGAGAGACTGATGCTCCGGAGTCCAGTCTTCAATATATGGAGA 1304
 QY 350 GlyValGlyThrLeuGluThrLeuPheTyrGlyLysGlyAspPheAspGluArgProGlu 369
 Db 1305 GGAGTGTATACCGGAGAGCTTTGATGATGAAAGAGAGATTCGATACCAACCAAGAG 1364
 QY 370 IleSerTyrGlyAspGlyAspGlyThrValAsnLeuValSerLeuLeuAlaLeuGlnSer 389
 Db 1365 ATTAGATATGAGATGAGATGGAGCGTATATTTGGGAGACTTACAGACTTTG----- 1418
 QY 390 LeuTyrLysGluGluLysAsnGlnTyrLeuLysValValLysIleAspGlyValSerHis 409
 Db 1419 ----- 1419
 QY 410 ThrSerIleLeuLysAsnGluValAlaLeuAsnGlnIleValGlyGlnIleThrSerIle 429
 Db 1464 ACATCTACTACTTAAAGACGAGATGCACTTAAAGAGATTATGAAGACAGATTTCATTTAT 1523
 QY 430 Asn 430
 |||

Db 1524 AAT 1526
 RESULT 8
 ID AAA49204
 AAA49204 standard; cDNA: 1500 BP.
 AC
 AA49204;
 DT 12-DEC-2000 (first entry)
 DE Corn putative lecithin:cholesterol acyltransferase gene #4.
 KW Corn; lecithin:cholesterol acyltransferase; phytosterol;
 XX phosphatidylcholine-sterol O-acyltransferase; heat shock; cold shock; ss.
 OS Zea mays.
 FH Key Location/Qualifiers
 FT CDS 24..1277
 FT /*tag= a
 FT /product= "lecithin:cholesterol acyltransferase"
 FT /EC_number= "2.3.1.43"
 XX
 PN WO200032791-A2.
 PD 08-JUN-2000.
 XX
 PF 02-DEC-1999; 99WO-US28586.
 PR 03-DEC-1998; 98US-0110782.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 PI Cahoon RE, Kinney AJ, Sakai H, Shen JB, Butler KH, Saylor JJ;
 DR WPI: 2000-412337/35.
 DR P-PSDB: AAB01209.
 XX
 PT Polynucleotide encoding plant lecithin:cholesterol acyltransferase
 PT enzyme useful for producing transgenic plants and for producing
 PT antibodies specific to which is useful for screening cDNA expression
 PT libraries -
 XX
 PS Claim 2; Page 41-42; 49pp; English.
 XX
 CC The present sequence is a putative coding sequence for a corn
 CC lecithin:cholesterol acyltransferase (also known as
 CC phosphatidylcholine-sterol O-acyltransferase). This enzyme is found
 CC associated with high-density lipoproteins and Apolipoprotein-AI and -D.
 CC The gene and protein can be used to produce transgenic plants which have
 CC increased lipid metabolism and membrane fluidity, and therefore increased
 CC resistance to heat and/or cold shock, to alter the content of phytosterol
 CC or lecithin in grains and to identify potential herbicides.
 CC
 SQ Sequence 1500 BP; 304 A; 469 C; 454 G; 271 T; 2 other:
 Alignment Scores:
 Pred. No.: 4,97e-88 Length: 1500
 Score: 913.50 Matches: 180
 Percent Similarity: 60.29% Conservative: 69
 Best Local Similarity: 43.58% Mismatches: 155
 Query Match: 39.16% Indels: 9
 Db: 21 Gaps: 4
 US-09-857-612a-14 (1-443) x AAA49204 (1-1500)
 QY 19 ThrValValValValMetLeuSerLeuLeuSerThrCysGlyAlaSerAsnLeuAspPro 38
 Db 60 ACGATCGTGTGTCCTCAAAATTTGCGAGCACGACGAGCGCCGACCGAAGACAGCTGCCGCC 119
 QY 39 LeuIleLeuIleProGlyAsnGlyLysGlnLeuGlnAlaArgLeuThrAsnGlnTyr 58
 Db 120 GTGCTGTGTGTGCGCCGGGTACGCCCAACAGAGCTCGACGGCGGCTCAGCGAGCTGTAC 179

```

Oy 59 LysProSerThrPheIleCysGluSerTrpTyProLeuIleLysLysAsnGlyTrp 78
    |||||::: ||| |||||
Db 180 CACCGCTCGCACCGCGCGC-----CGGCACAAAGGGAAGGCTG 221
Oy 79 PheArgLeuTrpPheAspSerSerValIleLeu---AlaProPheThrGlnCysPheAla 97
    |||||::: ||| |||||
Db 222 TTCGGCTCTACCTCAACACAGCGCGCTGAGAGCGCGCCGACCTGCGCTGTTCGCG 281
Oy 98 GluArgMetThrLeuIleIleTrpHisGlnGluLeuAspAspSerPheAsnThrProGluVal 117
    |||||::: ||| |||||
Db 282 GACGACATGCCACGCGCGCTACACGCGCGCTCCGACGACTACGCAACGCCACGCGGCGT 341
Oy 118 GluThrArgValProHisIleGlySerThrAsnSerLeuLeuIleAsnProAlaGlu 137
    |||||::: ||| |||||
Db 342 GAGACCGCGCTGCTTCTTCTGCGATCCACCGCGCTTCCGCTACCCGACCCGACGACGCG 401
Oy 138 LysHisIleThrGlyTrpMetAlaProLeuValAspSerLeuGlnLysLeuGlyTrpAla 157
    |||||::: ||| |||||
Db 402 AGAACTTCTCG---TACATGCAAGATTGCTCGCGCTGAGCGCGCTCGCTACCGC 458
Oy 158 AspGluGluThrLeuPheGluAlaProTyraAspPheArgTrpGlyLeuAlaIleGluGly 177
    |||||::: ||| |||||
Db 459 GACCGCGAAGACTGTTCGCGCGCGCTTCCGATCGACTTCGCTGACGCGCTCGCGCCGACGCG 518
Oy 178 HisProSerGlnValGlySerLysPheLeuLysAspLeuLysAsnLeuIleGluGluAla 197
    |||||::: ||| |||||
Db 519 CACCGCTCGAGGCTCGCAGCGCGCTTCTTCCGCGCGCTCAGAGGCTGTGAGAGAGGCGG 578
Oy 198 SerAsnSerAsnAsnGlyLysProValIleLeuLeuSerHisSerLeuGlyGlyLeuPhe 217
    |||||::: ||| |||||
Db 579 ACCCGGCTAACGAGGAGCGCGCTGACCATCGTGGCCACACATCGACGCGCGCACGCGT 638
Oy 218 ValLeuGlnLeuLeuAsnArgAsnProProSerTrpArgLysLysPheIleLysHisPhe 237
    |||||::: ||| |||||
Db 639 CGCGACCACTTCTACTGCGCGCGCGCTTCCGCTGCGCGAGCGCGCTTCCGCGCGCTTC 698
Oy 238 IleAlaLeuSerAlaProTrpGlyGlyAlaIleAspGluMetLysTrpPheAlaSerGly 257
    |||||::: ||| |||||
Db 699 GAGCGCGTTGCGCGCGCGCTGAGGAGCGTGTCTGCGATGTCATCATTCGCGCGCGG 758
Oy 258 AsnThrLeuGlyValProLeuValAspProLeuLeuValArgAspGluGlnArgSerSer 277
    |||||::: ||| |||||
Db 759 AACCATCTTCGCGCTGCTGCTGCGACCGCTGCTAACGGCGCATCGACGCGCTGCG 818
Oy 278 GluSerAsnLeuTrpLeuLeuProAsnProLysIlePheGlyProGluLysProIleVal 297
    |||||::: ||| |||||
Db 819 CAGACGAGCTCTGCGCGCTGCGCCCAACCCCAACGCACTTACAGACCGCGCGACGACTGCTG 878
Oy 298 IleThrProIleArgProTyrsSerAlaHisAspMetValAspPheLeuLysAspIleGly 317
    |||||::: ||| |||||
Db 879 ACCACACGAGGAGGACGACGACGCGCCACGACATGCGGACTCTCTACACGCGATCGG 938
Oy 318 PheProGluGlyValTyroTrpArgLysIleLeuProLeuIleGlyLysIleLys 337
    |||||::: ||| |||||
Db 939 CTAGGCGCGCGCAATTCGCTGCGACCACTCCGCGCTGCTCCCTGTCGCGGAGCGCGCA 998
Oy 338 AlaProGlnValProIleThrCysIleMetGlyThrGlyValGlyThrLeuGluThrLeu 357
    |||||::: ||| |||||
Db 999 TCTCCGCGCGCTGCGCTGCTGCTGCTGCGGCTTGGGCTGACACGCGCGGAGATGCTG 1058
Oy 358 PheTyrglyLysGlyAspPheAspGluArgProGluIleSerTyrglyAspGlyAspGly 377
    |||||::: ||| |||||
Db 1059 GCGTACCGCGGAGACGACTTCGACGTGACGCGCGATGATGTCATGAGAGACGCGCGG 1118
Oy 378 ThrValAsnLeuValSerLeuLeuAlaLeuGlnSerLeuTrpLysGluGluLysAsnGln 397
    |||||::: ||| |||||
Db 1119 CTGCTCAACCTGTGAGCTCTCTCTGCTGACCTCGCTGAGG---CTTCTCTACAGCT 1175
Oy 398 TyrLeuLysValValLysIleAspGlyValSerHisThrSerIleLeuLysAspGluVal 417
    |||||::: ||| |||||
Db 1176 TACTTTAGATCTCAAGGTGCGCAACGCTGTACACACGCGCGCTCTGCTGAGACGATGCT 1235

```

```

Oy 418 AlaLeuAsnGluIleValGlyGluIleThrSerIleAsn 430
    |||||::: ||| |||||
Db 1236 GCTCTCGCGCTTATCATTTAGCGCATCTACGCCCAT 1274
RESULT 9
AAA49205
ID AAA49205 standard; cDNA; 1660 BP.
AC AAA49205;
XX
XX 12-DEC-2000 (first entry)
DE
DE Corn putative lecithin:cholesterol acyltransferase gene #5.
XX
XX Corn: lecithin:cholesterol acyltransferase; phytosterol;
KW phosphatidylcholine-sterol O-acyltransferase; heat shock; cold shock; ss.
XX
XX Zea mays.
XX
XX Key Location/Qualifiers
FH CDS /tag= a
FT 77..1396 /product= "lecithin:cholesterol acyltransferase"
FT /EC_number= "2.3.1.43"
FT
PN WO200032791-A2.
XX
XX 08-JUN-2000.
XX
XX 02-DEC-1999; 99WO-US28586.
XX
XX 03-DEC-1998; 98US-0110782.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E. I.
XX
XX Cahoon RE, Kinney AJ, Sakai H, Shen JB, Butler KH, Saylor JJ;
XX WPI; 2000-412337/35.
XX DR P-PSDB; AAB01210.
XX
XX Polynucleotide encoding plant lecithin:cholesterol acyltransferase
XX enzyme useful for producing transgenic plants and for producing
XX antibodies specific to which is useful for screening cDNA expression
XX libraries.
XX
XX Claim 2; Page 43-44; 49pp; English.
XX
XX The present sequence is a putative coding sequence for a corn
XX lecithin:cholesterol acyltransferase (also known as
XX phosphatidylcholine-sterol O-acyltransferase). This enzyme is found
XX associated with high-density lipoproteins and Apolipoprotein-AI and -D.
XX The gene and protein can be used to produce transgenic plants which have
XX increased lipid metabolism and membrane fluidity, and therefore increased
XX resistance to heat and/or cold shock, to alter the content of phytosterol
XX or lecithin in grains and to identify potential herbicides.
XX
XX Sequence 1660 BP; 361 A; 475 C; 499 G; 325 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 2,936-69 Length: 1660
Score: 739..50 Matches: 171
Percent Similarity: 60.10% Conservative: 76
Best Local Similarity: 41.61% Mismatches: 141
Query Match: 31.70% Indels: 23
DB: 21 Gaps: 15
US-09-857-612A-14 (1-443) x AAA49205 (1-1660)
Oy 33 AlaSerAsnLeuAspProIleLeuIlePheProGluLysGlyAsnGlnLeuGluAla 52
    |||||::: ||| |||||
Db 179 GCCACACACTTCCACCCCATCTTCTGCTGAGTGGGCTGACGCTGACGACCTGAGGCA 238
Oy 53 ArgLeuThrAsnGlnTyrlLysProSerThrPheIleCysGluSerTrpTyProLeuIle 72

```



```

Percent Similarity: 54.55%      Conservative: 83
Best Local Similarity: 34.15%    Mismatches: 161
Query Match: 26.23%             Indels: 24
                                      Gaps: 10
US-09-857-612a-14 (1-443) x AAV68568 (1-1236)

QY 16 LeuThrValThrValValValMetLeuSerLeuLeuCysThrCysGlyAlaSerAsn 35
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 49 CTCGCTGTTACTCTGTTCTACTATGATGCTGGCAGACCTGACCTCCCGGCCACAGCT 108
QY 36 LeuAspProLeuLeuLeuLeuLeuProGlyLeuAsnGlyLeuGluAlaIleuArgLeuThr 55
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 109 CACCCCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 168
QY 56 AsnGlnTyrLysProSerThr-----PheIleCysGluSerTyrTyrProLeuIleLys 73
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 169 -----AAGCCAAAGGTGTACACTACCTTGTCTCC-----AAG 201
QY 74 LysLysAsnGlyTyrPheArgLeuThrPheAspSerSerValIleLeuAlaProPheThr 93
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 202 AAGACGGACAGCTACTTACACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 261
QY 94 GlnCysPheAlaGluArgMetThrLeuHisTyrHisGlnGluLeuAspAspTyrPheAsn 113
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 262 GACTGCTGATGATGACAAATATCAGCGCTGTTACACAGAACATCTCGGCCACCCAGTTT 321
QY 114 ThrProGlyValGluThrArgValProHisPheGlySerThrAsnSerLeuLeuThrLeu 133
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 322 CCCGATGCTGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 381
QY 134 AsnProAlaGluLeuLysHisIleThrGlyTyrMetAlaProLeuValAspSerLeuGlnLys 153
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 382 GACCCACCAAGAGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
QY 154 LeuGlyTyrAlaAspGlyGluThrLeuPheGlyAlaProTyrAspPheArgTyrGlyLeu 173
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 442 TGGGCTACACACAGCGGTGGAAGAGCTGAGAGTCTCCCTATGATGGGGGAGCC--- 498
QY 174 AlaIleGluGlyHisProSerGlnValGlySerLysPheLeuLysAspLeuLysAsnLeu 193
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 499 -----CCAAATGAAACGCGCCCTACTCTTGTGGCC---CTGGAGACAGATG 540
QY 194 IleGluGluAlaSerAsnSerAsnAsnGlyLysProValIleLeuLeuSerHisSerLeu 213
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 541 ATCGAGGAGATGTACCAATATGATGGGGC---CCGCTGCTGCTGCTGCTGCTGCTGCTGCT 597
QY 214 GlyGlyLeuPheValLeuGlnLeuLeuAsnArgAsnProProSerTyrPArgLysLysPhe 233
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 598 GGCACACGTGATCATCTCTACTTCTGCAGCGGAGCCACCAAGTCTGGAGAGACAAATAT 657
QY 234 IleLysHisPheIleAlaLeuSerAlaProTyrGlyAlaIleAspGluMetCtyrThr 253
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 658 ATCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
QY 254 PheAlaSerGlyAsnThrLeuGlyValProLeuValAspProLeuLeuValArgAspGlu 273
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 718 CTGGCTTACGAGACAAACATCGCATCCCGCTATGGGCCCTGAAGATCGGGAGACAG 777
QY 274 GlnArgSerSerGluSerAsnLeuThrLeuLeuProAsnProLysIlePheGlyProGln 293
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 778 CAGCATCTGCCGCTCTACACAGCTGGCTACGTCATCAACACCAACACTTGGTCACATGAA 837
QY 294 LysProIleValIleThrProIleArgProTyrSerAlaHisAspMetValAspPheLeu 313
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 838 AAGTATTGTTATACACACGACACTAATACAGCTCCGCGCTATACACCGGTTCTTC 897
QY 314 LysAspIleGlyPheProGluGlyValTyrProTyrGluThrArgIleLeuProLeuLeu 333
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 898 CGGCACATCGGTTTGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954
QY 334 GlyAsnIleLysAlaProGlnValArgProIleThrCysIleMetGlyThrGlyValGlyThr 353
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 955 GAAGCATGACGCCACCCGGGTGAGCTGCTGCTGTATAGCACTGCTGTCCACG 1014
QY 354 LeuGluThrLeuPheTyrGlyLysGlyAspPheAspGluArg---ProGluLeuSerTyr 372
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1015 CCAAACTCTTTTCTACTAC-----GAGACCTTCTCTGATGGGACCCCAAAATCTGCTTC 1068
QY 373 GlyAspGlyAspGlyTyrThrValAsnLeuValSerLeuLeuAlaLeuGlnSerLeuTyrLys 392
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1069 GCGCATGTGACGACGCGCATGAACTGAGACACGCTCTGCAATGCCAACCC---TGGCAG 1125
QY 393 GlnGluLysAsnGlnTyrLeuLysValValLysIleAspGlyValSerHisThrSerIle 412
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1126 ACCCCCAAGACACAGATATCATTCAGAGAGTGCAGCGGAAGCAGACATGTAGATG 1185
QY 413 LeuLysAspGluValAlaLeu 419
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1186 CTAGCCAAATGCCACCACTTG 1206

RESULT 12
AAV68566
ID AAV68566 standard; cDNA, 1236 BP.
XX
AC AAV68566;
XX
DT 16-FEB-1999 (first entry)
XX
DE Human heart lecithin-cholesterol acyltransferase-like gene 3.
XX
KW Human; heart; lecithin-cholesterol acyltransferase-like protein;
KW LCAT; agonist; arteriosclerosis; atherosclerosis; hyperlipidemia;
KW inflammatory disease; hypercalorism; obesity; hypertriglyceridemia;
KW malnutrition; abetalipoproteinemia; inflammatory disease;
KW Tangier disease; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1236
FT FT /*tag= a
FT FT /product= "LCAT protein"
XX
PN MO9846767-A1.
XX
PD 22-OCT-1998.
XX
PF 09-APR-1998; 98WO-JP01643.
XX
PR 22-JAN-1998; 98JP-0010289.
PR 11-APR-1997; 97JP-0093355.
PR 10-JUL-1997; 97JP-0184885.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Tanlyama Y;
XX
PI WPI: 1998-594583/50.
XX
DR P-PSDB: AAM80942.
XX
PT Isolated lecithin-cholesterol acyltransferase proteins - used to
PT develop products for treating e.g. atherosclerosis, obesity,
PT inflammatory diseases, senescence, diseases of the brain or renal
PT disorder.
XX
PS Example 1; Pages 144-145; 190pp; English.
XX
CC This is the nucleotide sequence encoding a novel human heart
CC lecithin-cholesterol acyltransferase-like protein, used in the method
CC of the invention. The novel proteins have LCAT-like activity. The
CC proteins and agonists can be used for treating or preventing e.g.
CC arteriosclerosis, atherosclerosis, hyperlipidemia, atherosclerosis
CC hypercalorism, obesity, hypertriglyceridemia, inflammatory diseases,
CC senescence, diseases of the brain, or renal disorder. Inhibitors of

```


QY 137 LeulysHisIleThrglyTyrrMetalProleuValAspSerLeuGlnLysLeuGlyTyr 156
 Db 411 AAAGCAGCGTGGCTCTATTCACACATGTCGAGGAGCTTGTGGCTGGGGCTAC 470
 QY 157 AlaspGlyGlnThrLeuPheGlyAlaProTyrTrpAspPheArgTyrGlyLeuAlaActu 176
 Db 471 ACACGGGGGTGAGATGCTCGAGGGGCTCTATGACTGCGCCGAGCC----- 518
 QY 177 GlyHisProSerGlnValGlySerLysPheLeuLysAspLeuLysAsnLeuIleGlnLu 196
 Db 519 -----CCAATGAAACGGGCGCTACTCTCTGGCC---CTCCGGAGATATATCGAGAG 569
 QY 197 AlaSerAsnSerAsnGlyLysProValIleLeuLeuSerHisSerLeuGlyLysLeu 216
 Db 570 ATGATACAGCTGTATGGGGG---CCCGTGTGCTGGTGGCCACAGATATGCGCAACTG 626
 QY 217 PheValLeuGlnLeuLeuAsnArgAsnProProSerTrpAlaGlyLysPheIleLysHis 236
 Db 627 TACACGCTCTACTTCTGACGCGCAGCGCAGCGCTGGAAGGAGCAACTATATCGGGCC 686
 QY 237 PheIleAlaLeuSerAlaProTyrGlyValAlaIleAspGlnMetTyrThrPheAlaSer 256
 Db 687 TTCCTGTACACTGGGTGGCCCTGGGGGGCGCTGGCCAAAGACCTGGCGCTCTGCTTCA 746
 QY 257 GlyAsnThrLeuGlyValProleuValAspProleuLeuValArgAspGlnIleArgSer 276
 Db 747 GGAGACAACAACGCGAGCCGATCGCGGCGCTGAAGATCCGGACACAGCGGCTCA 806
 QY 277 SerGlnSerAsnLeuThrLeuLeuProAsnProLysIlePheGlyProGlnLysProIle 296
 Db 807 GCTCTCTCCACACCTGCTGCTGCTCCCTACACATACATGCTGACCTGGAAGAGCTGTC 866
 QY 297 ValIleThrProIleArgProTyrSerAlaHisAspMetValAspPheLeuLysAspIle 316
 Db 867 GTGCAGACACCCCAATCACTACACACTGCGGGGCTACCCCAAGTTCTTCAGGACATC 926
 QY 317 GlyPheProGlnGly---ValTyrProTyrGlnThrArgIleLeuProleuIleGlyAsn 335
 Db 927 GCGTTTGAAATGGCTGGCTCATGCGGAGACAGCAAGNA-----GGGCTGGTGGAAACC 980
 QY 336 IleLysAlaProGlnValProIleThrcysIleMetGlyThrGlyValGlyThrLeuGln 355
 Db 981 ACGATGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1040
 QY 356 ThrLeuPheTyrGlyLysGlyAspPheAspGlnArg---ProGlnIleSerTyrGlyAsp 374
 Db 1041 TCTCTTACTAT-----GAGAGCTTCCCTGACCGTACCTAAATCTGCTTGTGTAC 1094
 QY 375 GlyAspGlyThrValAsnLeuValSerLeuLeuAlaLeuGlnSerLeuTrpLysGlnLu 394
 Db 1095 GGGCATGGTACTGTGAAGCTTAAGAGTCCCTGCAGTCCAGGCC---TGGCAGAGCCGC 1151
 QY 395 LysAsnGlnTyrLeuLysValValIleAspGlyValSerHisThsSerIleLeuLys 414
 Db 1152 CAGAGACACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1211
 QY 415 AspGlnValAlaLeu 419
 Db 1212 AACGCCACACCCCTG 1226
 RESULT 15
 ID AAA88516 standard; cDNA; 2680 BP.
 AC AAA88516;
 XX 22-JAN-2001 (first entry)
 DE Human PRO540 cDNA clone DNA44189-1322.
 KW PRO540; human; ocular disease; retinopathy; maculopathy; therapy;
 KW retinitis pigmentosa; macular degeneration; retinal detachment;
 KW retinal tear; macular hole; myopia; traumatic choriorretinopathy;

KW acute retinal necrosis syndrome; contusion; edema;
 KW retinal vision occlusion; vascular disease; retinal vasculitis;
 KW thrombocytopenic purpura; uveitis; retinal occlusion; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 21..1259
 FT sig_peptide 21..104
 FT mat_peptide 105..1256
 FT /*tag= c
 PN WO200053760-A2.
 XX 14-SEP-2000.
 PD 10-MAR-2000; 2000WO-US06319.
 XX 12-MAR-1999; 99US-0123957.
 PR (GETH) GENENTECH INC.
 XX Ferrara N, Goddard A, Gurney AL, Hebert C, Henzel WJ, Kabakoff RC;
 PI Klein RD, Kijavlin IJ, Kuo SS, La Fleur M, Wood WJ.
 XX WP1: 2000-587437/55.
 DR P-PSDB: AAB19579.
 XX
 PT Novel PRO polypeptides useful for preventing or rescuing retinal cells
 PT from injury caused by ocular diseases such as retinitis pigmentosa,
 PT retinopathy, retinal degenerative diseases, degenerative myopia,
 PT uveitis -
 XX
 XX Example 10: Fig 3; 140bp; English.
 PS
 XX The present sequence is that of cDNA clone DNA44189-1322 or UNQ341
 CC (ATCC 209699) isolated from a human foetal tissue cDNA library
 CC using primers and probes (see AAA88527-29) based on a consensus (see
 CC AAA88522) of sequences obtained from expressed sequence tag databases
 CC using secreted protein extracellular domains. The clone contains a
 CC single open reading frame encoding a 412-amino acid protein (see
 CC AAB19579) designated PRO540. The isolated cDNA can be used in the
 CC recombinant production of PRO540. The invention relates to the
 CC use of PRO polypeptides, including PRO540, to delay, prevent
 CC or rescue retinal cells such as retinal neurons selected from
 CC photoreceptors, retinal ganglion cells, displaced retinal ganglion
 CC cells, amacrine cells, displaced amacrine cells, horizontal and
 CC bipolar neurons, and supportive cells (including Mueller cells and
 CC pigment epithelial cells) from injury and degradation. The retinal
 CC cells are preferably photoreceptors and photoreceptor cell injury or
 CC death is caused by retinal injury, light or environmental trauma or
 CC by an ocular disease selected from retinitis pigmentosa, macular
 CC degeneration, including age-related, retinal detachment, retinal
 CC tears, retinopathy, retinal degenerative diseases, macular holes,
 CC degenerative myopia, acute retinal necrosis syndrome, traumatic
 CC choriorretinopathies or contusion such as Purtscher's retinopathy,
 CC edema, ischemic conditions such as central or branch retinal vision
 CC occlusion, collagen vascular diseases, thrombocytopenic purpura,
 CC uveitis, retinal vasculitis and occlusion associated with Fales
 CC disease and systemic lupus erythematosus (claimed).
 CC
 XX
 SQ Sequence 2680 BP; 518 A; 800 C; 782 G; 580 T; 0 other;
 Alignment Scores:
 Pred. No.: 4.98e-55 Length: 2680
 Score: 610.00 Matches: 141
 Percent Similarity: 54.57% Conservative: 80
 Best Local Similarity: 34.81% Mismatches: 156
 Query Match: 26.15% Indels: 28
 DB: 21 Gaps: 12

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2003, 14:06:02 ; Search time 50 Seconds
(Without alignments)
2717.156 Million cell updates/sec

Title: US-09-857-612a-14
Perfect score: 2333
Sequence: 1 MKKEDEGKIEVATLVTVY.....GEITSINSHAEGLSNLFSG 443

Scoring table:
BLOSUM62
Xgapop 10.0 , Ygapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09857612/r/unat_11022003_111819_29863/app_query.fasta.1.583
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=tni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=p2n -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09857612.@CGN_1_1_17.@unat_11022003_111819_29863 -NCPU=6 -ICPU=3
-NO_XLPEXT -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Issued_Patents_NA.*
- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	100.5	4.3	1533	1	US-08-483-232-23	Sequence 23, Appl
2	100.5	4.3	1533	2	US-08-485-938A-23	Sequence 23, Appl
3	100.5	4.3	1533	2	US-08-910-041-23	Sequence 23, Appl
4	100.5	4.3	1533	3	US-09-328-474-23	Sequence 23, Appl
5	100.5	4.3	1533	3	US-09-100-546-23	Sequence 23, Appl
6	100.5	4.3	1533	3	US-09-010-715-23	Sequence 23, Appl
7	100.5	4.3	1533	4	US-09-577-758-23	Sequence 23, Appl
8	99.5	4.3	4776	2	US-08-852-401-1	Sequence 1, Appl
9	98.5	4.2	2909	4	US-08-104-158-1	Sequence 1, Appl
10	97.5	4.2	1041	2	US-08-602-359A-24	Sequence 2, Appl
11	97	4.2	3048	2	US-08-313-200-2	Sequence 2, Appl
12	97	4.2	3048	5	PCT-US93-03837-2	Sequence 2, Appl

13	96	4.1	9510	4	US-09-453-702B-256	Sequence 256, App
14	94	4.0	1983	4	US-09-134-001C-513	Sequence 513, App
15	92.5	4.0	3128	3	US-08-716-449-1	Sequence 1, Appl
16	92	3.9	2191	1	US-08-318-905-22	Sequence 22, Appl
17	92	3.9	2191	1	US-08-483-232-22	Sequence 22, Appl
18	92	3.9	2191	1	US-08-483-140-22	Sequence 22, Appl
19	92	3.9	2191	2	US-08-485-938A-22	Sequence 22, Appl
20	92	3.9	2191	2	US-08-910-041-22	Sequence 22, Appl
21	92	3.9	2191	3	US-09-328-474-22	Sequence 22, Appl
22	92	3.9	2191	3	US-09-100-546-22	Sequence 22, Appl
23	92	3.9	2191	3	US-09-010-715-22	Sequence 22, Appl
24	92	3.9	2191	4	US-09-577-758-22	Sequence 22, Appl
25	91.5	3.9	3414	1	US-07-973-320-3	Sequence 3, Appl
26	90.5	3.9	5057	2	US-08-365-486A-12	Sequence 12, Appl
27	90.5	3.9	5057	2	US-08-860-342-12	Sequence 12, Appl
28	90.5	3.9	5108	1	US-07-642-002-1	Sequence 1, Appl
29	90	3.9	2381	2	US-08-318-826A-9	Sequence 9, Appl
30	90	3.9	2416	2	US-08-318-826A-8	Sequence 8, Appl
31	90	3.9	2416	4	US-09-334-489-1	Sequence 1, Appl
32	90	3.9	2416	4	US-09-334-489-2	Sequence 2, Appl
33	89.5	3.8	8268	1	US-08-375-709-10	Sequence 10, Appl
34	89.5	3.8	8268	1	US-08-752-929-10	Sequence 10, Appl
35	89.5	3.8	37895	1	US-08-375-709-1	Sequence 1, Appl
36	89.5	3.8	37895	1	US-08-752-929-1	Sequence 1, Appl
37	89.5	3.8	37895	3	US-09-090-793-1	Sequence 1, Appl
38	89	3.8	2444	3	US-08-906-791-1	Sequence 1, Appl
39	88	3.8	2400	6	5215909-13	Patent No. 5215909
40	87.5	3.8	1662	4	US-08-943-714-1	Sequence 1, Appl
41	87	3.7	1503	4	US-09-111-730-3	Sequence 3, Appl
42	87	3.7	1875	4	US-09-877-730-23	Sequence 23, Appl
43	86	3.7	2469	4	US-09-111-730-5	Sequence 5, Appl
44	86	3.7	1458	5	PCT-US94-03437-1	Sequence 1, Appl
45	86	3.7	1839	1	US-08-272-875-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
Sequence 23, Application US/08483232
Patent No. 5656431
GENERAL INFORMATION:
APPLICANT: Cousens, Lawrence S.
APPLICANT: Eberhardt, Christine D.
APPLICANT: Gray, Patrick W.
APPLICANT: Le Trong, Hai
APPLICANT: Tjoelker, Larry W.
APPLICANT: Wilder, Cheryl L.
TITLE OF INVENTION: Platelet-Activating Factor
TITLE OF INVENTION: Acetylhydrolase
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESSES:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,232
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,905
FILING DATE: 06-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/133,803
FILING DATE: 06-OCT-1993

ATTORNEY/AGENT INFORMATION:
 NAME: No. 565643land, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 27866/32689
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELE: 25-3658
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1533 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 62..1394
 US-08-483-232-23

Alignment Scores:
 Pred. No.: 0.0716 Length: 1533
 Score: 100.50 Matches: 108
 Percent Similarity: 33.52% Conservative: 69
 Best Local Similarity: 20.45% Mismatches: 153
 Query Match: 4.31% Indels: 198
 Gaps: 26

US-09-857-612a-14 (1-443) x US-08-483-232-23 (1-1533)

```

QY 9 LeuysIleGluValAlaThrLeuThrValAlaValAlaVal-MetLeuSerLeuLe 28
DB 37 CTGAAAGCGCAGCTCGAGATGTTACCGTCAATTCGATCCGCTTTCTGCCT 96
QY 28 uCystrHcyGlyAla-----SerAsnLeuAspProLeuIleLe 41
DB 97 CTGCACCTCCCTTGCACCTGTTATTCCTTTGACTGGCAAGACCTGAATCCAGTCCCTA 156
QY 41 uIleProGlyAsnGlyGly--AsnGlnLeuGluAlaArgLeuThrAsnGlnTyrLysPr 60
DB 157 TATTGAATCACCAGCATGGGTGCTAGTACAGTACAGCTGTGATGGCT----- 202
QY 60 oSerThrPheIleCysGlnSerTyrProLeuIleLysLysAsnGlyTyrPheAr 80
DB 203 -CCTCGAAACATTGGTCAATCTAATAATCCC-----AGAGGAATAGATCTTATTC 252
QY 80 g-----LeuTyrPheAspSerSerValIleLeuAlaProPheThrGlnCy 95
DB 253 CGTGGTGTACAGACTTGATGTTGAT----- 280
QY 95 sPheAlaGluArgMetThr-----LeuHisTyrHisGlnGluLeuAspAspTyrPh 112
DB 281 -TACACTAATAAGGACACTTCTGGCTTGTATATTCATCATCAAGATGATGATCACAC 339
QY 112 eAsnThrProGlyValGluThrArgVal-----Pr 122
DB 340 CCACACACCTTTGGATCCCAACAAAGAAATATTTTGGCTTAGTAATTTCTTGGAAAC 399
QY 122 oHis-----PheGlySerThrAsnSerLeuLeuLeu 132
DB 400 ACACAGTGGCTTGTGGGCAAAATATATGAGGCTTATCTTGGTCAATGACAACTCCGACAC 459
QY 132 rLeuAsnProArgLeuLysHisIleThrGlyTyrMetAlaProLeuValAspSerLeuGl 152
DB 460 CTGGAAATGCAATCTGAGG-----ACTGGGAAAAAATACCCACTAATATTTTTCCTCA 513
QY 152 nLysLeuGly----- 155
DB 514 TGGCTCTTGGAGCATTCAGAGCATTTATCTGCTATTGGCATTTGATCCGACACGG 573
QY 156 -----TyrAlaAspGlyGluThrLeuPheGlyAlaProTyrLys 168
DB 574 GTTATAGTCTGCTGTAGAACAGAGGATGGCTCTGCA-----TCCTGCACATACATA 627

```

```

QY 168 pPheArgTyrGlyLeuAlaAlaGluGlnHisProSerGlnValGlySerLys----- 185
DB 628 TTTCAGAGCAACAGCTGCTGTAGAA-----ATAGCGAACAAAGCTTGGCT 672
QY 186 -PheLeuLysAspLeuLysAsnLeuIleGluGluAlaSerAsnSerAsn----- 201
DB 673 CTATCTCAAGAACCTGAGACGAGAGAGGAGGATTCTCTTACGAAATAGACGATTAGC 732
QY 202 -----AsnGlyLys 204
DB 733 GCACAGCAAGAAAGAAATGTTCTCAACCTCTCACTTGATTGATTCGACATTGATCAGCGAG 792
QY 204 sProVal----- 206
DB 793 GCCAGTGAAGAAATGTACTAGATTAGATTGATGTGGAACAGCTGAGAGACTTATGTA 852
QY 207 -----IleLeuLeuSerHisSerLeuGlyGlyLeuPheValLeuGlnLeuLe 222
DB 853 TAGGATAAATAATAGCCATTATGGACATTCCTTGGTGGAGCCACAGATTATTCAGACTCT 912
QY 222 uAsnArgAsnProProSerTyrPheGlyLysPheIleLysHisPheIleAlaLeuSerAl 242
DB 913 TACTGAAGAC-----CAGAGATTCC--AGGTGGCATTTGCTCTGGATGC 954
QY 242 a-----ProTyrGlyGlyAlaIleAspGluMetTyrThrPheAlaSerGlyAsnTh 259
DB 955 ATGGATGTTTCCCGTGGGT-----CATGAAGATATAT-----TCCAG 990
QY 259 rLeuGlyValProLeuValAspProLeuLeuValArgAspGluGlnArgSerSerGluSe 279
DB 991 AATTCCTCAACCCCTC-----TTTTTTATCACTGGAGCATTCCTCAATACCCCTTC 1041
QY 279 rAsnLeuTyrProLeuProAsnProLysIlePheGlyProGlnLysProIleValIleTh 299
DB 1042 TAAAT-----ATCATTAAGATGAATAAATGCTTCTTACCTGATAGA----- 1081
QY 299 rProIleArgProTyrSerAlaHisAspMetValAspPheLeuLysAspIleGlyPheBr 319
DB 1081 ----- 1081
QY 319 oGluGlyValTyrProTyrGluThrArgIleLeuProLeuIleGlyAsnIleLysAlaPr 339
DB 1082 -----GAACGAAATAATGATTACAAATCAGGCGTTCGTCATCAGAA 1122
QY 339 oGlnValProIleThrCysIleMetGlyThrGlyValGlyThrLeuGluThrLeuPheTh 359
DB 1123 TTTTGTGACTTCACCTTTTGGCAGTACGAAATAATATGCTTACCTTACACTG----- 1177
QY 359 rGlyLysGlyAspPheAspGluArgProGluIleSerTyrGlyAspGlyLysPglyThrVa 379
DB 1178 ----AAAGAGACATGATTCATTCATGACCATTCACCTTACGACAAAGCTTCCTTACC 1233
QY 379 lAsnLeuValSerLeuLeuAlaLeuGln-----SerLeuTyrLys 392
DB 1234 GTTCTTTCAAAAACATTTTGGACTTCAGAAAGATTTTGTATGATGAGTGGGATTTCTAGTTA 1293
QY 392 sGluGluLysAsnGlnTyrLeuLysValValLysIleAspGlyValSerHisThrSerI 412
DB 1294 AGCGAAGATCCAAATCTTATTCAGAGGACCAACATTTAACACACACCAACCAAGCCAT 1353
QY 412 eLeuLysAspGluValAlaLeu 419
DB 1354 TCTGAGAGACTCCACAGAAATA 1375

```

RESULT 2
 US-08-485-938A-23
 ; Sequence 23, Application US/08485938A
 ; Patent No. 5847088
 ; GENERAL INFORMATION:
 ; APPLICANT: Cousens, Lawrence S.
 ; APPLICANT: Eberhardt, Christine D.
 ; APPLICANT: Gray, Patrick W.

APPLICANT: Le Trong, Hai
 APPLICANT: Tjoelker, Larry W.
 APPLICANT: Wilder, Cheryl L.
 TITLE OF INVENTION: Platelet-Activating Factor
 NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 City: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/485,938A
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/318,905
 FILING DATE: 06-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/133,803
 FILING DATE: 06-OCT-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5847088and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 27866/32792
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEX: 25-3658
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1533 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 62..1394
 US-08-485-938A-23
 Alignment Scores:
 Pred. No.: 0.0716 Length: 1533
 Score: 100.50 Matches: 108
 Percent Similarity: 33.52% Conservative: 69
 Best Local Similarity: 20.45% Mismatches: 153
 Query Match: 4.31% Indels: 198
 Gaps: 26
 US-09-857-612a-14 (1-443) x US-08-485-938A-23 (1-1533)
 QY 9 LeuylsilegluValAlaIthrLeuThValThrValValVal-MetLeuSerLeuLe 28
 Db 37 CTGAAGGAGGAGCTGACCTTCGAGATGTTACCTCCAAATTGCATGGCTTCTGCT 96
 QY 28 ucysThcysGlyAla-----SerAsnLeuAspProLeuIleLe 41
 Db 97 CTGACCTGCTGCTGCTGTTATTCCTTTGACTGCGAAGACCTGATCCAGTTGCCCTA 156
 QY 41 uileProGlyAsnGlyGly--AsnGlnLeuGluAlaArgLeuThrasnGlnTyrlsPr 60
 Db 157 TATTGATTCACACGACGAGCTAGATACAGCTCTGATGCT----- 202
 QY 60 oSerThPheIleCysGlnSerTyrProLeuIleLysLysLysAsnGlyTTPpPhear 80
 Db 203 -GGTCMAACATGTGTCATCTAAATCCCT-----AGAGGAATGATCTTATTTC 252

QY 80 g-----LeuTrpPheAspSerSerValIleLeuAlaProPheThrGlnCy 95
 Db 253 CGTCGGTGTACAGACTTGATTTGAT----- 280
 QY 95 sPheAlaGluArgMetThr-----LeuHisTyrHisGlnGluLeuAspAspTyrPh 112
 Db 281 -TACACTAATTAAGGACCTTCCTGCTTGATATTCATCCATCAAGATGATGATCACTC 339
 QY 112 easThrProGlyAlaGluThrArgVal-----Pr 122
 Db 340 CGACACCTTTGGATCCCAACAAAGAAATATTTTGGCTCTAGTAAATTTCTTGAAC 399
 QY 122 oHis-----PheGlySerThrAsnSerLeuLeuTy 132
 Db 400 ACACCTGCTTGTGGCAAAATTAATGGCTTATTTCTTCGTTCAATAGACACTCTGACGC 459
 QY 132 rLeuAsnProArgLeuLysHisIleThrGlyTyrMetAlaProLeuValAspSerLeuG1 152
 Db 460 CTGGAATGCATCTCGAGG-----ACTGGGAAATAATCCACTAATTTATTTTCTCA 513
 QY 152 nLysLeuGly----- 155
 Db 514 TGGTCTTGAGACATTCAGAGCATTTATTTCTGATTTGATTCAGATCCACAGG 573
 QY 156 -----TyrAlaAspGlyGluThrLeuPheGlyAlaProTyrAs 168
 Db 574 GTTATATAGTTCTGCTGTAGAACACAGGATGCTCTCA-----TCTTGACACTACTA 627
 QY 168 pPheArgTyrGlyLeuAlaAlaGluGlyHisProSerGlnValGlySerLys----- 185
 Db 628 TTTCAAGACACAGTCTGCTGTAGAA-----ATAGCAACAAAGTCTTGCT 672
 QY 186 -PheLeuLysAspLeuLysAsnLeuIleGluGluAlaSerAsnSerAsn----- 201
 Db 673 CTATCTCAGAACCTTGAGAGGAGAGAGAGAGTTCCTTACGAAATGAGCAGTTACG 732
 QY 202 -----AsnGlyLy 204
 Db 733 GCACGACAAAGAAATGTTCTCAAGCTCTCAGTTGATTTGACATTTGATCACGGGAG 792
 QY 204 sProVal----- 206
 Db 793 GCCAGTGACGATGTACTAGATTTAGAGTTGATGTGACACAGCTGAAGACTATTGA 852
 QY 207 -----IleLeuLeuSerHisSerLeuGlyGlyLeuPheValLeuGlnLeuLe 222
 Db 853 TAGGATTAATAATAGCATTTATTTGACATTTCTTTGGTGGACACAGTATTCAGACTCT 912
 QY 222 uAsnArgAsnProProSerTTPArgLysLysPheIleLysHisPheIleAlaLeuSerAl 242
 Db 913 TAGTGAAGAC-----CAGAGATTC--AGGTGGCATTGCTGTGGATGC 954
 QY 242 a-----ProTrpGlyGlyAlaIleAspGlnMetTyrThrPheAlaSerGlyAsnTh 259
 Db 955 ATGATGTTTCCCGTGGGT-----GATGAGATATAT-----TCCAG 990
 QY 259 rLeuGlyValProLeuValAspProLeuLeuValArgAspGlnArgSerSerGluSe 279
 Db 991 AATTCCTCAACCCCTC-----TTTTTATCAACTCGGACAGCATTCACATACCTTC 1041
 QY 279 rAsnLeuTrpLeuLeuProAsnProLysIlePheGlyProGlnLysProIleValIleTh 299
 Db 1042 TAAT-----ATCATAGAAATGAATAAAGCTCTTCACTCGATVGA----- 1081
 QY 299 rProIleArgProLysSerAlaHisAspMetValAspPheLeuLysAspIleGlyPhePr 319
 Db 1081 ----- 1081
 QY 319 oGluGlyValLysProLysArgIleLeuProLeuIleGlyAsnIleLysAlaPr 339
 Db 1082 -----GAACGAAATAATGATTACATTCAGGGGTGCGTCATCAGAA 1122

```

QY 339 oGlnValProIleThrCysIleMetGlyThrGlyValGlyThrLeuGluThrLeuPheTyr 359
DB 1123 TTTTGTGACTTCACCTTTTGCCACTAGCAAAATATTTGGCTACCTATTCCACTG----- 1177
QY 359 rGlyLysGlyAspPheAspGluArgProGluIleSerTyrGlyAspGlyAspGlyThrVal 379
DB 1178 ----AAAGAGACATGATTCATTCATAGCCATGACCTTAGCACAAGAAAGCTTCCTTAGC 1233
QY 379 LAsnLeuValSerLeuLeuAlaLeuGln-----SerLeuTyrPhe 392
DB 1234 GTTCTTACAAAAGATTTTGGACTTCAGAAAGATTTTGATTCATGGGATTTCTTACTGTGA 1293
QY 392 sGluGluLysAsnGlnTyrLeuLysValValLysIleAspGlyValSerHisThrSer11 412
DB 1294 AGGCGAAGATCACAATCTTATTCCAGGAGACACAACTTAACACAAACACACACACCAT 1353
QY 412 eLeuLysAspGluValAlaLeu 419
DB 1354 TCTGCAGAACTCCACAGAAATA 1375

RESULT 3
US-08-910-041-23
; Sequence 23, Application US/08910041
; Patent No. 5977308
;
; GENERAL INFORMATION:
; APPLICANT: Cousens, Lawrence S.
; APPLICANT: Eberhardt, Christine D.
; APPLICANT: Gray, Patrick W.
; APPLICANT: Le Trong, Hai
; APPLICANT: Thoecker, Larry W.
; APPLICANT: Wilder, Cheryl L.
; TITLE OF INVENTION: Platelet-Activating Factor
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,041
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,232
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/318,905
; FILING DATE: 06-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/133,803
; FILING DATE: 06-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELE: 25-3658
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1533 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 62..1394
; US-08-910-041-23

Alignment Scores:
Pred. No.: 0.0716 Length: 1533
Score: 100.50 Matches: 108
Percent Similarity: 33.52% Conservative: 69
Best Local Similarity: 20.45% Mismatches: 153
Query Match: 4.31% Indels: 198
DB: 2 Gaps: 26

US-09-857-612A-14 (1-443) x US-08-910-041-23 (1-1533)

QY 9 LeuLysIleGluValAlaIleThrValThrValValVal-MetLeuSerLeu 28
DB 37 CTGAAGCGCAGCTCAGCTTCGAGATGTTACCGTCCAATTCATGCGCTTTCTGCCT 96
QY 28 uCysThrCysGlyAla-----SerAsnLeuAspProLeuIle 41
DB 97 CTGCACCTCGCTTGACACTGCTGTTATCCCTTTGACTGGCAAGACGTGAATCCAGTTCCTTA 156
QY 41 uIleProGlyAsnGlyGly--AsnGlnLeuGluAlaArgLeuThrAsnGlnTyrLysPr 60
DB 157 TATTGAATCACCAGCATGGTGAGTAAATACAGCTGTGATGGCT----- 202
QY 60 oSerThrPheIleCysGlnSerTyrPyrProLeuIleLysLysAsnGlyTyrPheArg 80
DB 203 -CCTGCAAACTGGTCATCTTAATAATCCG-----AGAGAAATGATCTTATTC 252
QY 80 g-----LeuTyrPheAspSerSerValIleLeuAlaProPheThrGlnCys 95
DB 253 CCGCGTGTACAGACTTATGTTGAT----- 280
QY 95 sPheAlaGluArgMetThr-----LeuHisTyrHisGlnGluLeuAspAspTyrPh 112
DB 281 -TACACTAATAAGGCGACCTCTTGCGTTGTATATCATCATCAAGATGATGATCAGTC 339
QY 112 eAsnThrProGlyValGluThrArgVal-----Pr 122
DB 340 CGACACCCCTTGGATCCCAAGAAATATTTTGGCTTACTTAATTTCTTGGAAAC 399
QY 122 oHis-----PheGlySerThrAsnSerLeuLeuTyr 132
DB 400 ACACAGCGCTTGGGCAAAATATATGGCTTATTTCTGCTTAATGACAACTCTGCACG 459
QY 132 rLeuAsnProArgLeuLysHisIleThrGlyTyrMetAlaProLeuValAspSerLeuG 152
DB 460 CTGAATGCACATCTGAGG-----ACTGGGAAAAAATACCACTAATATATTTTTCGA 513
QY 152 nLysLeuGly----- 155
DB 514 TGGCTCTTGAGCATTCAGAGACATTTATTCGTATTTGGCATTTGATCCGACCGG 573
QY 156 -----TyrAlaAspGlyGluThrLeuPheGlyAlaProTyrAs 168
DB 574 GTTTATAGTTCCTGCTGTAAGAACACAGGATGCGTCTGCA-----TCTCGACACTACTA 627
QY 168 pPheArgTyrGlyLeuAlaIleGluGlyHisIleProSerGlnValGlySerLys----- 185
DB 628 TTTCAAGGACGAGCTGCTGTAGAA-----ATAGCAACAAGCTCTGGCT 672
QY 186 -PheLeuLysAspLeuLysAsnLeuIleGluGluAlaSerAsnSerAsn----- 201
DB 673 CTATCTCAGAAACCTCGAAGCGAGAGAGAGAGAGAGTTCCTTACGAAATGACACTTACG 732
QY 202 -----AsnGlyLys 204
DB 733 GCACAGCAAGAAAGATTTCTCAAGCTCTCAGTTGATTCTGGACATTTGATCAGCGAGG 792
QY 204 sProVal----- 206

```

```

Db 793 GCCAGTGCAGATGACTAGATTAGATTGATGTCGACAGACGTGAGACTCTATTGA 852
      :|||||
Qy 207 -----lleuleuSerHisSerleuGlyleuPheValleuGlnleuLe 222
      :|||||
Db 853 TAGGGATAAATAGCCATTATTGGACATCTTTGGTGGACACACGATTATTCACACT 912
      :|||||
Qy 222 uasnArgasnProProSerTrpArgLysPheIleLysHisPheIleValleuSerAl 242
      :|||||
Db 913 TAGTGAGAC-----CAGAGATTC---AGGTGTGGCATTTGCTTGATGC 954
      :|||||
Qy 242 a-----ProTrpGlyGlyAlaIleAspGluMetTrpThrPheAlaSerGlyAsnTh 259
      :|||||
Db 955 ATGATGTTTCCCGTGGT-----GATGAGGTATAT-----TCCAG 990
      :|||||
Qy 259 rleuGlyValProleuValAspProleuLeuValArgAspGluGlnArgSerSerGluSe 279
      :|||||
Db 991 AATTCCTCACCCCTC-----TTTTTATCAACTCGGACGATTCACATACCCCTTC 1041
      :|||||
Qy 279 rAsnLeuTrpLeuProAsnProLysIlePheGlyProGlnLysProIleValIleTh 299
      :|||||
Db 1042 TAAT-----ATCATAGAAATGAAAAATGCTTCTTACCTGATAGA----- 1081
      :|||||
Qy 299 rProIleArgProTyrSerAlaHisAspMetValAspPheLeuLysAspIleGlyPhePr 319
      :|||||
Db 1081 ----- 1081
      :|||||
Qy 319 oGluGlyValIleProTyrGluThrArgIleleuProleuIleGlyAsnIleLysAlaPr 339
      :|||||
Db 1082 -----GAACGAAAAATGATTACATCAGGGCTTGCCATCAGAA 1122
      :|||||
Qy 339 oGlnValProIleThrCysIleMetGlyThrGlyValGlyThrleuGluThrleuPheTy 359
      :|||||
Db 1123 TTTTGTTACTTCACTTTTGGCACTAGCAAAATATGCTTACTTACCTGATAGA----- 1177
      :|||||
Qy 359 rGlyLysGlyAspPheAspGluArgProGluIleSerTyrGlyAspGlyAspGlyThrVa 379
      :|||||
Db 1178 -----AAAGAGACATCGATTCACATGATAGCCATCAGCTTGACAAACAGCTTCTTACG 1233
      :|||||
Qy 379 lAsnLeuValSerleuLeuAlaLeuGln-----SerleuTrpLys 392
      :|||||
Db 1234 GTTCTTACAAAACATTTAGACTTCAGAAAGATTTGATCAGTGGATTTCTTATTGATGCA 1293
      :|||||
Qy 392 sGluGluLysAsnGlnTyrleuLysValIleLysPheGlyValSerHisThrSerI 412
      :|||||
Db 1294 AGCGCAGATACAACTTTATTCAGGAGCAACATTTAACACCAACCAACCAACCAAT 1353
      :|||||
Qy 412 eLeuLysAspGluValAlaLeu 419
      :|||||
Db 1354 TCTGCAGAACTCCACAGGAATA 1375
      :|||||

```

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/328,474
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/483,232
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,905
FILING DATE: 06-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/133,803
FILING DATE: 06-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rln-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3658
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1533 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: CDS
LOCATION: 62..1394
US-09-328-474-23
Alignment Scores:
Score: 0.0716 Length: 1533
Percent Similarity: 100.50 Matches: 108
Best Local Similarity: 33.52% Conservative: 69
Query Match: 20.45% Mismatches: 153
DB: 4.31% Indels: 198
Gaps: 26
US-09-857-612a-14 (1-443) x US-09-328-474-23 (1-1533)
Qy 9 LeuLysIleGluValAlaThrLeuThrValIleValIleValIleMetLeuSerLeuLe 28
      :|||||
Db 37 CTGAAACGCGACCTCAGCTTCGAGATGTTACCGTCCAAATTCAGATCGCCTTTCTGCCT 96
      :|||||
Qy 28 uCysThrCysGlyAla-----SerAsnLeuAspProleuIleLe 41
      :|||||
Db 97 CTGACACTGCTGTTGACACTGTTATTCCTTTGACTGCAAGACCTGAAATCCAGTGCCTA 156
      :|||||
Qy 41 uLleProGlyAsnGlyGly---AsnGlnleuGlnAlaArgleuThrAsnGlnTyrLysPr 60
      :|||||
Db 157 TATTGAATCACCAGCAGTGGTGCAGTAAGTACAGCTGTGATGCT----- 202
      :|||||
Qy 60 oSerThrPheIleCysGluSerTrpTyrProleuIleLysLysLysAsnGlyTrpPheAr 80
      :|||||
Db 203 -GCTGCAACATTTGGTCATCTAAATCC-----AGAGGAATGATCTTATTC 252
      :|||||
Qy 80 g-----LeuTrpPheAspSerSerValIleleuAlaProPheThrGlnCy 95
      :|||||
Db 253 CCTCGTGTGTACAGACTTGATTTGAT----- 280
      :|||||
Qy 95 sPheAlaGluArgMetThr-----LeuHisTyrHisGlnGlnleuLeuAspTyrPh 112
      :|||||
Db 281 -TACACTAATTAAGGACCTTCTTGCGTTGATTAATTCATCATCAGATGATGATCAGTGC 339
      :|||||
Qy 112 eAsnThrProGlyValGluThrArgVal-----Pr 122
      :|||||

```


DB: 3 Gaps: 26

US-09-857-612a-14 (1-443) x US-09-100-546-23 (1-1533)

Qy 9 LeuylsllleglValAlaThLeuThrValThrValValValVal-MetLeuSerLeuLe 28
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 37 CTGAAGGCGAGCTCGAGATGTACCGTCCAAATGTGACATGGCTTTCTGCT 96
 Qy 28 uCysThrCysGlyAla-----SerLeuAspProLeuLeLe 41
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 97 CTGCACCTGGCTTGCACGTGTTATCTTTCCTTTCAGTCGCAAGACCTGATTCAGTTGCC 156
 Qy 41 uileProGlyAsnGlyLeu---AsnGlnLeuGluAlaArgLeuThrsAsnGlnTyLysPr 60
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 157 TATGATTCACACAGCATGGGCTAGATACAGCTCTGATGGCT----- 202
 Qy 60 oSerThrPheLeuGlySerLysTrpTyrrProLeuIleLysLysAsnGlyTrpPheAr 80
 : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 203 -GCTGCAACATGTGTAATCTAAATCCCC-----AGAGCAAAATGGATCTTATTC 252
 Qy 80 g-----LeuTrpPheAspSerSerValIleLeuAlaProPheThrGlnCy 95
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 253 CGTGGTGTACACACTTGATGTTGAT----- 280
 Qy 95 sPheAlaGluArgMetThr-----LeuHisTyrrHisGlnGluLeuAspArgTyrrPh 112
 : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 281 -TACATAATTAAGGACCTTCTGCTTGTATTCATTCATCAGATGATGATCAGTAC 339
 Qy 112 eAsnThrProGlyAlaGluThrArgVal-----Pr 122
 : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 340 CGACACCCCTTGGATGCCCAAAAGAAATATTTTGGGCTGTAGTAATTTCTTGGAAC 399
 Qy 122 oHis-----PheGlySerThrsAsnSerLeuLeuTy 132
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 400 ACATGGCTTGTGGGCAAAATATATGGCTTATCTTCGCTTCATGACATCCCTCGCAGC 459
 Qy 132 rLeuAsnProArgLeuLysHisIleThrGlyTyrrMetAlaProLeuValAspSerLeuG 152
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 460 CTGGAATGCACATCTGAGG-----ACTGGGAAATAATCCACTAATATTTTTCACA 513
 Qy 152 nLysLeuGly----- 155
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 514 TGGCTTGGAGCATTCAGAGCATTTATTCGTATTCGATTCATTCGATCCACGCG 573
 Qy 156 -----TyrAlaAspGlyGluThrLeuPheGlyAlaProTyrrAs 168
 : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 574 GTTATAGTCTGCTCTAGAACACAGGATGGCTTCGA-----TCCTGACATCTACTA 627
 Qy 168 pPheArgTyrrGlyLeuAlaAlaGluGlyHisProSerGlnValGlySerLys----- 185
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 628 TTTCAGGACACAGTCTGTAGAA-----ATAGCAACAAGTCTTGCT 672
 Qy 186 -PheLeuLysAspLeuLysAsnLeuIleGlnAlaLaserAsnSerAsn----- 201
 : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 673 CTATCTCAGAACCTTGAGGAGAGAGAGAGAGAGTTCCTTACGAAATAGAGCAGTTACG 732
 Qy 202 -----AsnGlyLys 204
 : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 733 GCAACGAGCAAAAGAAATGTTCTCAAGCTCTCAGTTTGTGACATTCATCACGGGAG 792
 Qy 204 sProVal----- 206
 : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 793 GCCAGTGCAGATGTACTAGATTAGATTGATGTGGAACAGCTCAAGCTCATTTGA 852
 Qy 207 -----IleLeuLeuSerHisSerLeuGlyGlyLeuPheValLeuGlnLeuLe 222
 : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 853 TAGGATAAATATACCATTTATTTGACATTTCTTTGGTGAGCCACAGTTATTCAGACTCT 912
 Qy 222 uAsnArgAsnProProSerTrpArgLysLysPheIleLysHisPheIleAlaLeuSerAl 242
 : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 913 TAGTGAAGAC-----CAGAGATTC--AGGTGGCATTCCTGCGATGC 954
 Qy 242 a-----ProTrpLysGlyAlaAlaLeuAspGluMetTyrrThrPheAlaSerGlyAsnTh 259

Db 955 ATGATGTTTCCCGTGGT-----GATGAGCTATAT-----TCCAG 990
 Qy 259 rLeuGlyValProLeuValAspProLeuLeuValArgspGluArgSerSerGluSe 279
 : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 991 AATTCCTCAACCCCTC-----TTTTTATACACTCGGAGACATTCATACCTTC 1041
 Qy 279 rAsnLeuTrpLeuLeuProAsnProLysIlePheGlyProGlnLysProIleValIleTh 299
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 1042 TAAT-----ATCATTAAGAAAGAAAATGCTCTTACTGTATAGA----- 1081
 Qy 299 rProIleArgProTyrrSerAlaHisAspMetValAspLeuLysAspIleGlyPhePr 319
 Db 1081 ----- 1081
 Qy 319 oGluGlyValTyrrProTyrrGluThrArgIleLeuProLeuIleGlyAsnIleLysAlaPr 339
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 1082 -----GAACGAAAATGATTAATCAATCAGGGGTGCTGCATCAGAA 1122
 Qy 339 oGlnValProIleThrCysIleMetGlyThrGlyValGlyThrLeuGluThrLeuPheTy 359
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 1123 TTTTGTGACTTCACTTTTGGCAGTACGAAATAATATGCTATTCACACTG----- 1177
 Qy 359 rGlyLysGlyAspPheAspGluArgProGluIleSerTyrrGlyAspGlyAspGlyThrVa 379
 : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 1178 -----AAAGAGACATCGATTCCAATGTAGCCATCAGCCTTAGCAACAAAGCTTCCTTAGC 1233
 Qy 379 rAsnLeuValSerLeuLeuAlaLeuGln-----SerLeuTrpLys 392
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 1234 GTTCTTACAAAACATTTAGAGCTTCAGAAAAGATTTTGAATCAGTGGATTCCTTATGTTGA 1293
 Qy 392 sGluGluLysAsnGlnTyrrLeuLysValValLysIleAspGlyValSerHisThSerIl 412
 : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 1294 AGCGCAAGATCAGATCTTATTCAGGAGCAGACATTAACACACCAACCAACCAAGCAT 1353
 Qy 412 eLeuLysAspGluValAlaLeu 419
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 1354 TCTGCAGAACCTCCACAGGAATA 1375

RESULT 6
 US-09-010-715-23
 : Sequence 23, Application US/09010715
 : Patent No. 6146625
 : GENERAL INFORMATION:
 : APPLICANT: Cousens, Lawrence S.
 : APPLICANT: Eberhardt, Christine D.
 : APPLICANT: Gray, Patrick W.
 : APPLICANT: Le Trong, Hai
 : APPLICANT: Tjoelker, Larry W.
 : APPLICANT: Wilder, Cheryl L.
 : TITLE OF INVENTION: Platelet-Activating Factor
 : TITLE OF INVENTION: Acetylhydrolase
 : NUMBER OF SEQUENCES: 30
 : CORRESPONDENCE ADDRESS:
 : ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
 : STREET: 6300 Sears Tower, 233 South Wacker Drive
 : CITY: Chicago
 : STATE: Illinois
 : COUNTRY: United States of America
 : ZIP: 60606-6402
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/010,715
 : FILING DATE:
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/318,905
 : FILING DATE: 06-OCT-1994
 : PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/133,803
 FILING DATE: 06-OCT-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: NO. 6146625and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 27866/32793
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-0448
 TELEFAX: 25-3658
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1533 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 62..1394
 US-09-010-715-23

Alignment Scores:
 Pred. No.: 0.0716 Length: 1533
 Score: 100.50 Matches: 108
 Percent Similarity: 33.52% Conservative: 69
 Best Local Similarity: 20.45% Mismatches: 153
 Query Match: 4.31% Indels: 198
 DB: Gaps: 26

US-09-857-612a-14 (1-443) x US-09-010-715-23 (1-1533)

QY 9 LeuLysIleGluValAlaThrLeuThrValThrValValVal-MetLeuSerLeu 28
 DB 37 CTGAAACGGCAGCTAGCTCGAGATGTACCGTCAATATGCAATGCGCTTTCGCT 96
 QY 28 uCyThrCysGlyAla-----SerAsnLeuAspProLeuIleLe 41
 DB 97 CTGCACCTGCTTCGCTGCTTTATTCCTTTGACGTGCAAGACCTGAATCCAGTTCCTTA 156
 QY 41 uLleProGlyAsnGlyGly--AsnGlnLeuGluAlaArgLeuThrAsnGlnTyrIle 60
 DB 157 TATTGAATCACACAGCATGGGTGATAGATACAGCTCTGATGGCT----- 202
 QY 60 oSerThrPheIleCysGluSerTyrTrpProLeuIleLysLysAsnGlyTrpPhe 80
 DB 203 -CCTGCAACATTGGTCATCTAATATCC-----AGAGGAATAGATCTTATTC 252
 QY 80 g-----LeuTrpPheAspSerSerValIleLeuAlaIleProPheThrGlnC 95
 DB 253 CCGTGGTGTACAGACTTGATGTTGAT----- 280
 QY 95 sPheAlaGluArgMetThr-----LeuHisTyrHisGlnGluLeuAspAspTyr 112
 DB 281 -TACACTAATAGGCGACCTTCTGCGTTGTATTCATCCATGTCAGATGATGATCACC 339
 QY 112 eAsnThrProGlyValGluThrArgVal-----Pr 122
 DB 340 GCACACCCCTTGGGCAAAATATATGCGCTTATTTAGTAATTTCTTGCAAC 399
 QY 122 oHis-----PheGlySerThrAsnSerLeuLeu 132
 DB 400 ACACAGCGCTTGGGCAAAATATATGCGCTTATTTAGTAATTTCTTGCAAC 459
 QY 132 rLeuAsnProArgLeuLysIleThrGlyTyrMetAlaProLeuValAspSerLeu 152
 DB 460 CTGGAATGCACATCTGAGG-----ACTGGGAAAAAATACCCACTAATATTTTTCGA 513
 QY 152 nLysLeuGly----- 155
 DB 514 TGGCTTGGAGCATTCAGACATTTATTCCTATTTGGCATTTGATCCACGCG 573
 QY 156 -----TyrAlaAspGlyGluThrLeuPheGlyAlaProTyrAs 168

DB 574 GTTATAGTTCCTCTGTAGAACACAGCGATGGCTCTGCA-----TCTCGACATAC 627
 QY 168 pPheArgTyrGlyLeuAlaAlaGluGlnHisProSerGlnValGlySerLys----- 185
 DB 628 TTTCAGAGACCACTCTGCTGTAGA-----ATAGGCAACAAAGTCTTGCT 672
 QY 186 -PheLeuLysAspLeuLysAsnLeuIleGluGluAlaSerAsnSerAsn----- 201
 DB 673 CTATCTCAGAACCCCTGACAGAGAGAGAGGTTTCTTTACGAATAGACAGATTAG 732
 QY 202 -----AsnGly 204
 DB 733 GCACGACCAAGAAATGTTCTCAACCTCACTTGAATTTGACATTGATCAGCGAG 792
 QY 204 sProVal----- 206
 DB 793 GCCAGTGACCAATGTACTGATTTAGAGTTTGATGTGGAACAGCTGAAGACTCTATTGA 852
 QY 207 -----IleLeuLeuSerHisSerLeuGlyGlyLeuPheValLeuGlnLeu 222
 DB 853 TAGGATAAATATAGCCATTATTGACATCTTTGGTGGAGCAGCATTTATTCAGACTCT 912
 QY 222 uAsnArgAsnProProSerTyrArgLysIlePheIleLysHisPheIleAlaLeuSerAl 242
 DB 913 TAGTGAGAGC-----CAGAGATTC--AGGTGTGCGATTCCTGATGC 954
 QY 242 a-----ProTyrGlyGlyAlaIleAspGluMetTyrThrPheAlaSerGlyAsp 259
 DB 955 ATGGATGTTTCCCGTGGT-----CATGAGATAT-----TCCAG 990
 QY 259 rLeuGlyValProLeuValAspProLeuLeuValArgAspGluGlnArgSerSerLeu 279
 DB 991 AATTCCTCAACCCCTC-----TTTTTATCAACCTGGAACAGATTCCATACCTTC 1041
 QY 279 rAsnLeuThrPheLeuProAsnProLysIlePheGlyProGlnLysProIleValIle 299
 DB 1042 TAAT-----ATCATAGAAATGAAAAAATGCTTCTTACCTGATAGA----- 1081
 QY 299 rProIleArgProTyrSerAlaHisAspMetValAspPheLeuLysAspIleGlyPhe 319
 DB 1081 ----- 1081
 QY 319 oGluGlyValTyrProTyrGluThrArgIleLeuProLeuIleGlyAsnIleLysAla 339
 DB 1082 -----GAACGAAATATGATACAAATCAGGGTTCGGTCCATCAGAA 1122
 QY 339 oGlnValProIleThrCysIleMetGlyThrGlyValGlyThrLeuGluThrLeuPhe 359
 DB 1123 TTTTGTGACTTCACCTTTTGCACATGAAATATTTGGCTACCTATTCACACTG----- 1177
 QY 359 rGlyLysGlyAspPheAspGluArgProGluIleSerTyrGlyAspGlyAspGlyThr 379
 DB 1178 -----AAAGGACACATGATTCATCCATGACCATTCACCAACAAAGCTCTTAC 1233
 QY 379 lAsnLeuValSerLeuLeuAlaLeuGln-----SerLeuTyr 392
 DB 1234 GTTCTTACAAAAAATTTAGACTTCAGAAAGATTTTGATCGTGGAGATTCTTGATGA 1293
 QY 392 sGluGluLysAsnGlnTyrLeuLysValValLysIleAspGlyValSerHisThrSer 412
 DB 1294 AGCGCAAGATCACAATCTTATTCAGAGGACCAACATTTAACACAAACCAACCAAGCAT 1353
 QY 412 eLeuLysAspGluValAlaLeu 419
 DB 1354 TCTGAGAACTCCACAGAAATA 1375

RESULT 7
 US-09-577-758-23
 ; Sequence 23, Application US/09577758
 ; Patent No. 6203790
 ; GENERAL INFORMATION:
 ; APPLICANT: Cousens, Lawrence S.

```

1 APPLICANT: Eberhardt, Christine D.
2 APPLICANT: Gray, Patrick W.
3 APPLICANT: Le Trong, Hai
4 APPLICANT: Tjoelker, Larry W.
5 APPLICANT: Wilder, Cheryl L.
6 TITLE OF INVENTION: Platelet-Activating Factor
7 TITLE OF INVENTION: Acetylhydrolase
8 NUMBER OF SEQUENCES: 30
9 CORRESPONDENCE ADDRESS:
10 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
11 STREET: 6300 Sears Tower, 233 South Wacker Drive
12 CITY: Chicago
13 STATE: Illinois
14 COUNTRY: United States of America
15 ZIP: 60606-6402
16 COMPUTER READABLE FORM:
17 MEDIUM TYPE: Floppy disk
18 OPERATING SYSTEM: IBM PC compatible
19 SOFTWARE: Patentin Release #1.0, Version #1.25
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/09/577,758
22 FILING DATE:
23 CLASSIFICATION:
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: 09/010,715
26 FILING DATE:
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER: US 08/133,803
29 FILING DATE: 06-OCT-1993
30 ATTORNEY/AGENT INFORMATION:
31 NAME: No. 6203790and, Greta E.
32 REGISTRATION NUMBER: 35,302
33 REFERENCE/DOCKET NUMBER: 27866/32793
34 TELECOMMUNICATION INFORMATION:
35 TELEPHONE: (312) 474-6300
36 TELEFAX: (312) 474-0448
37 TELEX: 25-3658
38 INFORMATION FOR SEQ ID NO: 23:
39 SEQUENCE CHARACTERISTICS:
40 LENGTH: 1533 base pairs
41 TYPE: nucleic acid
42 STRANDEDNESS: single
43 TOPOLOGY: linear
44 MOLECULE TYPE: protein
45 FEATURE:
46 NAME/KEY: CDS
47 LOCATION: 62..1394
48 US-09-577-758-23
49
50 Alignment Scores:
51 Pred. No.: 0.0716 Length: 1533
52 Score: 100.50 Matches: 108
53 Percent Similarity: 33.52% Conservative: 69
54 Best Local Similarity: 20.45% Mismatches: 153
55 Query Match: 4.31% Indels: 198
56 DB: 4 Gaps: 26
57
58 US-09-857-612A-14 (1-443) x US-09-577-758-23 (1-1533)
59
60 9 leuylsliegIvalAlaThrLeuThrValThrValAlaValAl-MetleuSerLeuLe 28
61 ||||| :||||| :|||
62 Db 37 CTGAAACGGCAGCTCACGTTCGGAGATGTTTACCCTCAAAATTGCATCCGCTTTTCGCCCT 96
63
64 Oy 28 ucSThrTrcSclgYala-----SerAsnLeuaspProleuilele 41
65 ||||| ||| :||| :|||
66 Db 97 CTCGACCGCTTGACACTGCTTTATCTCTTTTGACTGGCAGACCTCATCAGTTGCCCTA 156
67
68 Oy 41 uIlleProclYasnlglyl--AasnGlneUglAlaArgLeuThrasnGlnTrlySpr 60
69 ||| :||| :|||
70 Db 157 TATTGAATCACCACAGATGGGTCACTAAGATACAAAGCTCTGTATGGCT----- 202
71
72 Oy 60 oSerThrpheIIecysgluSerTrpTyrrProleuIllelYslYslasnglYrrPhear 80

```

[illegible]

Db	1082	-----GACGAAAAATGATTACAACTCAGGGGTGGCTGCATCAGAA	1122
QY	339	ocInValProlIePhrCysIleMeTolYhrGlyValAlGlyThrIeGluThrLeuPheTy	359
Db	1123	TTTTGTGTGACTCTCACTTTTCCCACTGACCAAAATAATATGGCTACATTCACACTG	1177
QY	359	rgLYLysGlyAspPheAspLysrGrProGluIleSerTYrgLYAspGlyAspGlyThrVa	379
Db	1178	-----AAAGGAGACATCGATTCCAAATGATAGCACACCTTACCAACAAAGCTTCCTAGC	1233
QY	379	lAsnLeuValSerLeuLeuLeuAlaLeuGln-----SerLeuThrPly	392
Db	1234	GTTTTCACAAAACAACTTTAGAGACTTCAGAAAGACTTTTGATGCTAGTGGACTTCCTTAAGTTGA	1293
QY	392	sgLGIuLYLysAsnGlnTYrLeuLYsValLYsIleAspGlyValSerHisThrSerI	412
Db	1294	AGCGGAGATCACAACTTATTCACAGGAGCAACAACTTACACAAACCAACCAACGACCAT	1353
QY	412	eLeuLYsAspGlyValAlaLeu	419
Db	1354	TCTCGAGACCTCCACAGGAAT	1375

```

RESULT 8
US-08-852-401-1
Sequence 1, Application US/08852401
Patent No. 5976836
GENERAL INFORMATION:
APPLICANT: Weber, J. Mark
APPLICANT: Hessler, Paul E.
APPLICANT: Larsen, Peter E.
APPLICANT: Luu, B. Minh
TITLE OF INVENTION: Methods and Compositions for Enhancing
TITLE OF INVENTION: Erythromycin Production
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rockey, Milnamow & Katz, Ltd.
STREET: 2 Prudential Plaza, 180 N. Stetson, Suite
STREET: 4700
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,401
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mueller, Lisa L.
REGISTRATION NUMBER: 38,978
REFERENCE/DOCKET NUMBER: FEB2159POO30US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-616-5400
TELEFAX: 312-616-5460
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4776 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-852-401-1

Alignment Scores:
Pred. No.: 0.584 Length: 4776
Score: 99.50 Matches: 77
Percent Similarity: 32.25% Conservative: 52
Percent Local Similarity: 19.25% Mismatches: 131
Query Match: 4.26% Indels: 141

```

Alignment Scores:	
Pred. No.:	0.584
Score:	99.50
Percent Similarity:	3.25%
Best Local Similarity:	19.25%
Query Match:	4.26%
Length:	4776
Matches:	77
Conservative:	52
Mismatches:	131
Indels:	141

DB:	2	Gaps:	16
US-09-857-612A-14 (1-443)	x	US-08-852-401-1 (1-4776)	
QY	82	trpPheaspSerSerValIleLeuAlaProPheThrGlnCysPheAlaGlnArgMetThr	101
Db	1910	TGGTGGGGCGACGGGGGGTGTACAGAGTCTACGTCGGACGCTTGGCGCAC	1960
QY	102	LeuHisTyrHisGlnGluLeuAspAspTyrPheAsnThrProGlyIleValGluThrArgVal	121
Db	1961	-----GCCGACGGCGACGGGATTCGGCGACCTGGCGGGGTGGCGCCAAAGCTG	2008
QY	122	ProHisPheGlySerThrAsn-----SerLeuLeuTyrLeuAsnProArgLeuHisIle	140
Db	2009	CCGTAACCTGGTGGACCTGGGTGGACGGCGGTGGCTACACCCG-----	2053
QY	141	ThrGlyTyrMetAlaProLeuValAsp-----SerLeuGlnIleGluGlyTyrAlaAspGly	159
Db	2054	---TTCATACCGGTGGCGAATGGCGCAAGGGGTACGACGTGGCGACACTACGACAGTC	2110
QY	160	GluThrLeuPheGlyAlaProTyrAspPheArgTyrGlyLeuAlaAlaGluGlyHisPro	179
Db	2111	GAACCGATGTTCGGCACG-----	2128
QY	180	SerIleValGlySerIysPheLeuIysAspLeuIysAsnLeuIleGluIleValSerIle	199
Db	2129	-----CTCGACAGCTTCGACGACCTCTGGCGGGCG-----	2161
QY	200	SerAsnAsnGlyIysProValIleLeuLeuSerHisSerLeuGlyGlyLeuPheValLeu	219
Db	2162	-----CACTGCTGTGGCTGAAGTAGTACTTC	2188
QY	220	GlnLeuLeuAsnArgAsnProProSerTyrArgIysIysPheIleHisIleAla	239
Db	2189	GACGTCTGGTCCCAACACACACTCCGACGGCGCACCCCTGGTGGCGGAG-----GCG	2239
QY	240	LeuSerIleProIrrProGlyIleAlaIleAspGluMetTyrThrPheAlaSerGly	257
Db	2240	CTGGAGCGCGGGCGGGCGACCGCGCGGAGAGGCTACTCTGTCCGACGGCGCGCC	2299
QY	258	-----AsnThrLeuGlyValProLeuValAsp	266
Db	2300	GAGAGCGGGGAGCTGCCGCCCAACGACTGGAGATCATCTTGGGGGTCCGGC-GTGAC	2358
QY	267	ProLeuLeuValArgAspGluArgSerSerGluSerAsnLeuTyrLeuLeuProAsn	286
Db	2359	CCG-----	2361
QY	287	ProIysIlePheGlyProGlnIysProIleValIleThrProIleArgProIysSerIle	306
Db	2362	-----CGTCCCGACGGCCAGGTGATCTCGACCTGCTGTTCGCCCGACACACC	2409
QY	307	His-----Asp-MetValAspPheLeuIysAspIleGlyPhe-	318
Db	2410	CGACCTGAACCTGGCGCAACCCGAGATTCGGCGCGAGATTCCGCCAAGTGTCTGGATTCTG	2469
QY	319	---ProGluGlyValTyrTyrProTyrGlyIleThrArgIleLeuProLeuIleGlyAsnIleIy	337
Db	2470	GCTGGACCGGGGGGTGACGAGGCTTC-----CGGATCGACGTCCGCCACGAGATGATCAA	2523
QY	337	GAlaProGlnValProIleThrCysIleMetGlyThrGlyValGlyThrLeuGluThrIle	357
Db	2524	GCACCCCGACCTGCC-----GACACGGGGCGCCACGACGATCTCCCT	2568
QY	357	uPheTyrGlyIysGlyAspPheAspIleArgProGluIleSerTyrGlyAspGlyAspGly	377
Db	2569	GCCTC-----GGCCGGGCGAGCTGCCCTACTTTCGACAGAGACGA	2607
QY	377	YThrValAsnLeuVal-----SerLeuLe	385
Db	2608	GGTGGACGGGATCTACCGGGGAGTGGGGCGAGCTGCTGACACTCTACAGAGGGCGCCGAT	2667
QY	385	uAlaLeuGlnSerLeuTyrIysGluIleIysAsnGlnTyrIleIysValIleIysIleAs	405

OY	377	YTHValAsnLeuVal-----SerLeuLe	385
		::: :::	
Db	2608	GGTGCACGCGATCTACCGGAGTGGCGCGAGCTGCTGCACCTCTACGAGGCGCGCCGGAT	2667
OY	385	ValAlaLeuGlnSerLeuTrpLysClnGlnLysasnGlnTrpLeuLysValValLysIleAs	405

```

Db 2668 CGGGGTGCGCGAGCGTGGGCCCGGACCAAGTCAAGCGCTGGCCGCTACGTGGCCCGCA 2727
Qy 405 pglValSerHisThr----- 410
Db 2728 CGAGCTGCACCGAGCGTTCAACATGGCGCTCTGGAGTCCGCCGTGGCCGACGCGTT 2787
Qy 411 -----SerIleuLysAspGluValAlaLeuAsnGluIleValGlyGluIleThrSer 428
Db 2788 CCGGCGGTGTCATGCACGACGCTCGCGGCGCCACGACCGCGTCCGGGGCCACGACGACC 2845

RESULT 9
US-08-104-158-1
Sequence 1, Application US/08104158
Patent No. 6215042
GENERAL INFORMATION:
APPLICANT: Willmitzer, Lothar
APPLICANT: Sonnewald, Uwe
APPLICANT: Kossmann, Jens
APPLICANT: Mueller-Roeber, Bernd
APPLICANT: Visser, Richard Gerardus Franciscus
APPLICANT: Jacobsen, Evert
TITLE OF INVENTION: PLASMIDS CONTAINING DNA-SEQUENCES THAT
TITLE OF INVENTION: CAUSE CHANGES IN THE CARBOHYDRATE CONCENTRATION AND THE
TITLE OF INVENTION: CARBOHYDRATE COMPOSITION IN PLANTS, AS WELL AS PLANT CELLS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostroienk, Faber, Gerb & Soffen
STREET: 1180 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/104,158
FILING DATE: 13-AUG-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP PCT/EP92/00302
FILING DATE: 11-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 41 04 782.6
FILING DATE: 13-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meilman, Edward A.
REGISTRATION NUMBER: 24,735
REFERENCE/DOCKET NUMBER: FA-1996 PCT (951-91)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-382-0700
TELEFAX: 212-382-0888
TELEX: 236925
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2909 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Solanum tuberosum cv. Desiree
STRAIN: Desiree
DEVELOPMENTAL STAGE: growing tuber
TISSUE TYPE: tuber
CELL TYPE: total tuber
IMMEDIATE SOURCE:

```

```

; LIBRARY: cDNA of total tuber mRNA in pUC 19 (Hinc II)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1699
; OTHER INFORMATION: /note="for Branching enzyme I
; OTHER INFORMATION: (partial) truncated protein; 97,11 % identity to
; OTHER INFORMATION: active potato branching enzyme"
US-08-104-158-1

Alignment Scores:
Pred. No.: 0.344 Length: 2909
Score: 98.50 Matches: 93
Percent Similarity: 32.91% Conservative: 63
Best Local Similarity: 19.62% Mismatches: 151
Query Match: 4.22% Indels: 167
DB: 4 Gaps: 25

US-09-857-612A-14 (1-443) x US-08-104-158-1 (1-2909)
Qy 32 GLVALASerAsnLeuAspProIleuIleProGlyAsnGlyAsnGluIleu 51
Db 152 GGCTCCCTAAATTGGATCCA-----ACTTTGGA 181
Qy 52 AlaArgLeuThrAsn-----GlnTyrLysProSerThrPheIleCys 65
Db 182 CCTATCTAGATCACTTCAGACACAGAAATGAAGATATGTGATGACAAATGCTATT 241
Qy 66 GluSerTyrPyr---ProIleuIleLysLysAsnGlyTTPheArgLeuThrPhe--- 83
Db 242 GAAAAATATGAGGAGCCCTTGAGGAATTCCTCAAGGTATTTAAAAATTGGATTCAAC 301
Qy 84 -----AspSerSerValIleu-----AlaProPheThrGln----- 94
Db 302 AGGAGATGTGTTGCATGCTATGCTGAATGGCTCTCCTGCTCAGAGAAACAGAACTT 361
Qy 95 -----CysPheAlaGluArgMetThrLeuHisTyrHisGlnGluLeuAspAspTyr--- 111
Db 362 ATGGCGCATTTCAATGATGAGAGAGCGTTCTAACCAATGATGAGAAAGACAGAGTTGGT 421
Qy 112 -----PheAsnThrProGlyValGluThrArg-----ValProHisPheGlySer 126
Db 422 GTTGGAGTATATGAAATTCCTGATGTGATGACGTAAGCAGTCACTCCACAC----- 472
Qy 127 ThrAsnSerLeuTyrLeuAsnProArgLeuIleHisIleThrGlyTyrMetAlaPro 146
Db 473 ---AACTCCAGAGTTAACTTT---CGTTTCACACATGTAATGAGTGTGGTAACT 523
Qy 147 LeuValAspSerLeuGlnLysLeuGlyTyrAlaAspGlyGluThrLeuPheGlyAlaPro 166
Db 524 CGTATCCCTGCTTGATTAAGATGATGCCACGTCGACAGCGCC---ACAAAGTTTGCACACCA 580
Qy 167 TyrAsp-----PheArgTyrGlyLeu 173
Db 581 TATGATGCTGTCTACTGGACCCACCACTTCAGAAAGTACCACTCAATATACCTCGC 640
Qy 174 AlaAla-----GluGlnHisProSerGlnValGlySer--- 184
Db 641 CCTCCCAACCCCGAGCCCAAGAAATTCAGACAGACATGTCGGCATACAGCTCTAG 700
Qy 185 -----LysPheLeuLysAspLeuLysAsnLeuIleGluIleuAla 197
Db 701 CCAGGTAAATTGCTATCGATGAGTGTGCAGATGATGTTTAACTCGGATTT---AAGCA 757
Qy 198 SerAsnSerAsnAsnGlyLysProValIleLeuLeuSerHisSerLeuGlyGlyLeuPhe 217
Db 758 AATAACTATTAATACGTCCAGTGTGATGCCATTAAGGAACATTTCTTACTATGATCAATT 817
Qy 218 ValLeuGlnLeuLeuAsnArgAsnProProSerTTPArgLysLysPheIleLysHisPhe 237
Db 818 GGATATCATGTTACAAAC-----TTT 838
Qy 238 IleAlaLeuSerAlaProTyrGlyGlyAlaIleAspGluMetTyrThrPheAlaSerGly 257

```



```

QY 385 euAlaLeuGlnSerLeuThrPlyGluGluLysAsnGln----- 397
Db 1283 CGGCACTGACACGCTGTGCTCGCGAGCACAACCGCTGCGCGCGCTCAAGGCC 1342
QY 398 -----TyrLeuLysValValLysIleAspGlyValS 408
Db 1343 TCATATGCGCACTGAGCGCGAGCCGCTGTATCCAGAGGCGCGAGGTGTGCGCGCTC 1402
QY 408 erHisThrSerIle-----LeuLysAspGluVal-----AlaLeuA 420
Db 1403 TGCACGATCATCATCCCTGAGGATTACTCCCAAGATCTCGGAGCCGAGCCCTTCC 1462
QY 420 snGluIleValGlyGluIleThrSerIleAsnSerHisAlaGluLeuGlyLeuSerAsnL 440
Db 1463 AGCAGTACGTCGCTCCCTATGAGAGCTATGATGATCACCAGCCCAACCCACCTGTCCACAG 1522
QY 440 eupHeSer 442
Db 1523 TGTTCTCC 1530

RESULT 12
PCT-US93-03837-2
; Sequence 2: Application PC/TUS9303837
; GENERAL INFORMATION:
; APPLICANT: Baker, Jr., James R.
; APPLICANT: Koenig, Ronald J.
; TITLE OF INVENTION: Thyroid Peroxidase Epitopic Regions
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/03837
; FILING DATE: 19930422
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lewak, Anna M.
; REGISTRATION NUMBER: 33006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (313) 641-1600
; TELEFAX: (313) 641-0270
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3048 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; DEVELOPMENTAL STAGE: Mature
; TISSUE TYPE: Thyroid gland (from patients with
; TISSUE TYPE: Grave's disease)
; IMMEDIATE SOURCE:
; CLONE: pHPPO-2.8
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 2pter-q11
; PUBLICATION INFORMATION:
; AUTHORS: Kimura, S.
; AUTHORS: Kotani, T.
; AUTHORS: McBride, O. W.

```

```

;
; AUTHORS: Umeki, K.
; AUTHORS: Nakayama, T.
; AUTHORS: Ohkaki, S.
; AUTHORS: Hirai, K.
; TITLE: Human thyroid peroxidase: Complete cDNA and
; TITLE: protein sequence, chromosome mapping, and
; TITLE: identification of two alternatively spliced mRNAs
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
; VOLUME: 84
; PAGES: 5555-5559
; DATE: 1987
; RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 3048
PCT-US93-03837-2

Alignment Scores:
Pred. No.: 0.557 Length: 3048
Score: 97.00 Matches: 98
Percent Similarity: 33.78 Conservative: 52
Best Local Similarity: 22.07% Mismatches: 146
Query Match: 4.16% Indels: 149
DB: Gaps: 24

US-09-857-612A-14 (1-443) x PCT-US93-03837-2 (1-3048)
QY 84 AspSerSerValIleLeuAlaProPheThrGlnCysPheAlaGluArgMetThrLeuHis 103
Db 393 GAAACTCAACATCATCAGCATCC--AACGATGCTTATGAGAGATCTGCTGAGCAT 449
QY 104 Tyr-----HisGlnGluLeuAspAspPyrPheAsnThr 114
Db 450 CATTCGAACATGTCGTGATGTCCTTACATGATGCCCCCAAAATGCCCCAATGCTTG 509
QY 115 ProGlyValGluThrArgValProHisPheGlySerThrAsnSerLeuThrLeuAsn 134
Db 510 CCTGGCGAACAATACAGGCCCATCC-----AGAGCTTGCACACACAGAGA 557
QY 135 Pro-----ArgLeuLysHisIleThrGlyTyrMetAlaProLeu-ValAspSe 150
Db 558 CCACCCGATGAGGCGGCTCCCAACAGCGCCCTGGCAGCATGCTCCCTCCAGTC----- 612
QY 150 rleuGlnLysleuGlyTyrAlaAspGlyGluThrIlePheGlyAlaProTyrAspPheAr 170
Db 613 -----TATGAGAGCGC-----TTCACTACAGCCCGAGGCTGGAA 647
QY 170 gTyrGlyLeuAlaAlaGluGlyHisProSerGlnValGlySerLysPheLeuLysAspLe 190
Db 648 CCGCGCTCTTGTACAGCGGTCCCACTGCCCGGTCGGAGAGTACACAAGATGT 707
QY 190 uLysAsnLeuIleGluGluAlaSerAsnSerAsnAsnGlyLysProValIleLeuLeuSe 210
Db 708 CATTCAGTTTCAATGATGAGTGTGCACAGATGATGACCGCTATTCTGACCTCTGTATG-- 765
QY 210 rHisSerLeuGlyGlyLeuPheValLeuLnuLeuAsnAArgAsnProProSerTyrPr 230
Db 766 -----GCATGG-- 771
QY 230 glyLysPheIleLysHisPheIleAlaLeuSer-----AlaProTr 244
Db 772 -GGACATATCATCGACACAGACATCGCTTCACACACAGACACAGCAAGACGTGCCTT 830
QY 244 pGlyGlyAlaIleAspGluMetTyrThrPheAlaSerGlyAsnThrLeuGlyValProLe 264
Db 831 CGGGGAGGGGCTACTGCCAGATGACTGTGAGAAACCAACCA----- 876
QY 264 uValAspProLeuValArgAspGluGlnArgSerSerGluSerAsn-----LeuTr 282
Db 877 -TGTTTTCCCATACAACTCCGAGAGGCGCGCGCGCGGACCGCTGTCTGCGC 935
QY 282 pleuLeuProAsnProLysIlePhe-----GlyProGlnLysProIl 296
Db 936 -CTTCTACG-----CTCTTGGCGCGCTGGCGGACCGGAGACCAAGCGCGCTCT 985
QY 296 eValIleThrProIle-----ArgProTyrSerAlaHisAspMetValAspPheLeuLy 314

```



```

Db      8724 AGTAATGACCGGAACATCAAGATTTTCATCTAAATTAATTCAGGATTTATTTTAAAC 8783
QY      240 ---LeuSerIaIaProTrrp-glyGlyAlaIaIaLeuAspGluMetTyr-ThrPheIaSerGly-- 257
Db      8784 AACCTCAAAAACATCATGTCGTGGCTGGCTGTTGTTAATTTTAAAGATTTTATTAATACATG 8843
QY      258 -----AsnThrIeuGlyValProLeu----- 264
Db      8844 GTGCAATATATATCTTACTTATTAACAGTTCCTATTGATTAATTAATTCACACATAG 8903
QY      265 -----ValAspProIeuLeuValArg---AspGluGlnArgSerSerGluSerAsnIe 281
Db      8904 AATTAATATTAACCCATGCGGATGAAGATACAGAAACAGTCAAGTCAAGTCAAGTCAAGT 8963
QY      281 uTrpLeuLeuProAsnPro-----LysIlePheGlyProGlnLysProIleValI 298
Db      8964 GTTTTTCCTTACCTTACCTTACCTTAAATAAATGATGTTTATTCCTGCAATTAATCCACTTACTC 9023
QY      298 eThrProIleArgProTyrSerIaIaHisAspMetValAspPheLeuLysAspIleGlyPh 318
Db      9024 TAACATCATTCCTTCTTATTAATTCACATGAATTAATTTGATTTGATGA-----GCATTA 9077
QY      318 eProGluGlyValTyrProTyrGluThrArgIleLeuProLeuIleGlyAsnIleLysAl 338
Db      9078 TAGAAATCCAAAGCTATTCGAT-----TTACAGAACAAAT----- 9111
QY      338 aProGlnValProIleThrCysIleMetGlyThrGlyValGlyThrLeuGluThrLeuPh 358
Db      9112 -----ATTATCAATAGATATATACATACACTGTTTCT----- 9141
QY      358 eTyrGlyLysGlyAspPheAspGluArgProGluIleSerTyrGlyAspGlyLysAl 378
Db      9142 -----GAGTATGCTGCAAAAACACGCTAT 9164
QY      378 rValAsnIeu---ValSerIleuLeuAlaIeuGlnSer-----LeuTrrpLysGluGly 935
Db      9165 TGCCATATATGCTATTAATATATACGATATACGATATGCTGTTTGTGCAAAAGAAAG 9224
QY      395 sAsnGlnTyrLeuLysValValLysIleAspGlyValSerHis-----Th 410
Db      9225 TGTGCGTGTAGAGATACATATAAATTTGCTATTAATGACCATATGACTTTAGCTTTAA 9284
QY      410 rSerIleLeuLysAsp 415
Db      9285 TTCATTTAGTAAGAT 9300

RESULT 14
US-09-134-001C-513
; Sequence 513, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 513
; LENGTH: 1983
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-513

Alignment Scores:
Pred. No.: 0.626 Length: 1983
Score: 94.00 Matches: 90
Percent Similarity: 36.41% Conservative: 56

```

```

Best Local Similarity: 22.44% Mismatches: 104
Query Match: 4.03% Indels: 151
DB: 4 Gaps: 26

US-09-857-612A-14 (1-443) x US-09-134-001C-513 (1-1983)

QY      71 LeuIleLysLysLysAsnGlyTrrpPheArgLeuTrrpPheAspSerSerValIleLeuAla 90
Db      502 CTGTAAAGAAAGAACGATATTCCTCAACATT----- 534
QY      91 ProPheThrGlnCysPheAlaGluArgMetThrLeuHisTyrHisGlnGluLeuAspAsp 110
Db      535 -----AATTAATATACAGACCGCTTA---TTGAATTTTACATGAATAATCCAGAC 582
QY      111 Tyr-----PheAsnThrProGlyVal 117
Db      583 TTTATCAACACCACTATAGAAAAATGAATGATTAATTAATCTTTATCAACACAGGTTTA 642
QY      118 Gly-----ThrArgValProHisPheGly 125
Db      643 GAAGATTTAGCAAGTATCAGTACATCATTCGATGGGGGTACGCTGACCATCT----- 696
QY      126 SerThrAsnSerLeuLeuTyrIleAsnProArgLeuLysHisIleThr---GlyTyrMet 144
Db      697 -----AATCCT-----AAACATGTTGTATACGTGTGGATTT 726
QY      145 AlaProLeuValAspSerLeuGlnLysLeuGlyTyr---AlaAspGlyGluThrLeuPhe 163
Db      727 GATGCACTTGTTATATATTTCTTCATTAAGTTATCTGATCATGATGAACATTAATTT 786
QY      164 GlyAlaProTyrAspPheArgTyrGlyLeuAlaIaIaGluGlyHisProSerGlnValGly 183
Db      787 AAT-----AAATATTGGCCAGCA----- 804
QY      184 SerLysPheLeuLysAspLeuLysAsnLeuIleGluGlnAlaSerAsnSerAsnGly 203
Db      805 -----GACATACACTTGATGCTTAAGAA----- 828
QY      204 LysProValIleLeuLeuSerHisSerLeuGlyGlyLeuPheValLeuGlnLeuAsn 223
Db      829 -----ATTGTACGTTCCACTATTT-----ATTATGCCAATATTTGTAATGGCG 873
QY      224 ArgAsnProProSerTrrpArgLysLysPheIleLysHisPheIleAlaLeuSerAlaPro 243
Db      874 TTGGATTTACCACTTCCTAAAGATTTTTCACACGCGTGGATTTTAATGAAG----- 927
QY      244 TrpGlyGlyAlaIleAspGluMetTyrThrPheAlaSerGlyAsnThrLeuGlyValPro 263
Db      928 -----GATGTAATAATGAGTAATCTAAAGATTAAT----- 957
QY      264 LeuValAspProLeuLeuValArgAspGluGlnArgSerSerGluSerAsnLeuTrrpLeu 283
Db      958 GTGCTATGATCCCTAATATTAATTAATGATGCTTATGCTTATGATGCGACGATTAATCTACTTA 1017
QY      284 -----LeuProAsnProLysIlePheGlyProGlnLysProIleValIleThrPro 300
Db      1018 ATGCGGTGATTAACG-----TTTGGTTCGATGGC-----GTATTTACACCG 1059
QY      301 IleArgProTyrSerAlaHisAspMetValAspPheLeuLysAspIleGly---PheProG 320
Db      1060 -----GAAGCCTTGTGAAGAACAAATTAAGATCTTTCGATGAATGATTAATCTAGTG 1116
QY      320 IuGlyValTyrProTyrGluThrArgIleLeuProLeuIleGlyLysAlaIaProG 340
Db      1117 AATCGTACTATCTATATGATTAACAATTAATTTCCACG----- 1154
QY      340 IuValProIleThrCysIle-MetGlyThrGlyValGlyThrLeuGluThrLeuPheTyr 359
Db      1155 -----CGAATTACCTGATACCAAGGTCCA----- 1179
QY      360 GlyLysGlyAspPheAspGluArgProGluIle-----SerTyr 372
Db      1180 ---AAACATGAATTTGATGAAGAAATGGAAGGATGGCGCTTGAACATGTTAAATCATTC 1236

```

```

Oy 373 GLYSPGGLYASPGLYTHrVAlAsnLLeuValSerLeu-----LeuAlaLeuGlnSer 389
    |||
Db 1237 AATGAT-----AATATGCAAGATTTCACAAATTTCTGTGGTTATATCAAC 128
    :|||||
Oy 390 LeuTrpLys-----GluGluLysAsnGlnTyLeuLysValValLysIleAspGlyVal 407
    :|||||
Db 1282 GTATGCAAAATTATTAAGTCGTACAAACAATAAT-----ATTGATGAAACT 132
    :|||||
Oy 408 SerHisThrSerIleLeuLysAspGluValAlaLeuAsnGluIleValGlyGluIle 426
    :|||||
Db 1327 CAACCTTGCGGTTCTTGCAAAAAGATGAA---AATCAACGTGAGATGCTGTGTAATGTA 1380
    :|||
RESULT 15
US-08-716-449-1
Sequence 1, Application US/08716449
Patent No. 6103893
GENERAL INFORMATION:
APPLICANT: National Starch and Chemical Investment
APPLICANT: Holding Corporation
TITLE OF INVENTION: Method for Producing Altered Starch
TITLE OF INVENTION: from Potato Plants
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scott & Aylen
STREET: Box 194, 24th Floor, Toronto-Dominion Bank Tower
CITY: Toronto
STATE: Canada M5K 1H6
COUNTRY: Canada
COMPUTER READABLE FORM:
MEDIUM TYPE: 1MB 1.44 MB High Density Diskette
COMPUTER: COMPAQ Deskpro 590 (IBM PC compatible)
OPERATING SYSTEM: WINDOWS 95
SOFTWARE: Word 7.0 for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/716,449
FILING DATE: FILED CONCURRENTLY HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/00634
FILING DATE: 22.03.95
ATTORNEY/AGENT INFORMATION:
NAME: Anita E.Nador
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 1552
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 368-7240
TELEFAX: (416) 363-7246
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3128 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: cDNA encoding starch branching enzyme
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE: clone 1.2.1 and E2
ORGANISM: Solanum tuberosum
STRAIN: cv desiree
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE: mature tuber
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY: lambda zap tuber cDNA
CLONE: 1.2.1 and E2
POSITION IN GENOME:

```

:	CHROMOSOME/SEGMENT:
:	MAP POSITION:
:	UNITS:
:	NAME/KEY: open reading frame
:	LOCATION: 44-2788
:	IDENTIFICATION METHOD: lone ORF with homology to other starch
:	IDENTIFICATION METHOD: branching enzymes
:	OTHER INFORMATION: complements KV832 E. coli glycogen
:	OTHER INFORMATION: branching enzyme mutant
:	US-08-716-449-1
Alignment Scores:	
Pred. NO.:	1.96
Score:	92.50
Percent Similarity:	32.48%
Best Local Similarity:	18.68%
Query Match:	3.96%
DB:	3
US-09-857-612A-14 (1-443) x US-08-716-449-1 (1-3128)	
QY	35 AsnAsnSPProLeuIleLeuIleProGlyAsnGlyIleuGlLeuGlAlaArgLeu 54
Db	341 AATTGGATCCA-----ACITTGGAACTTAATA 370
QY	55 ThrAsn-----GlnTyrlYsProSerThrPheIleCysGluSerTrp 68
Db	371 GATCACTTCAGACACAGACAATGAAGAATGTGATCGACAGAAAATGCATTTGAAAAATAT 430
QY	69 Tyr---ProleuIleLysLysLysAsnGlyTrpPheArgLeuTrpHe-----asp 84
Db	431 GAGGAGCCCTTGAGCAATTTGGCTCAAGGTTATTTAAATTTGGATTCAACAGGAMAAT 490
QY	85 SerSerValIleLeuAlaProPheTrnGlnCysPheAlaGluArgMetHrileu----- 102
Db	491 GGTTGCATAGCTCATGCTGATAGTAGGCGCTCGCTGCCTCAGAAAGCACAAACTTATGGCAT 550
QY	103 -----HisTyrlHisGlnGluLeuAspAspTyrr----- 111
Db	551 TTCAATGGATGGAACGGSTTCTAACACATGATGGAAGAGACAGTTGGTGTGGAGT 610
QY	112 PheaSnThrProGlyValGluThrArg-----ValProHisPheGlySerThraSser 129
Db	611 ATTGAATTTCTCGTGTGTTCAGATAAGCCAGTCAATTCACAC-----AACTCC 658
QY	130 LeuLeuTyrlLeuAsnProAlaGlyLeuLysHisIleThrGlyTyrlMetAlaProLeuValasp 149
Db	659 AGAGTPTAGTTT-----CGTTTCAAACATGTAATGGAGTGGGTATGATCGTATGCCCT 712
QY	150 SerLeuGlnLysLeuGlyTyrlAlaAspGlyGluThrLeuPheGlyValaProTyrlasP-- 168
Db	713 GCTTGGATAAAGTATGCCACTGCAGACGCC--ACAAATTTGCACACACATATATATGCT 769
QY	169 -----PheArgTyrlGlyLeuAlala--- 175
Db	770 GTTACTGAGCCACCACCACTTCAGAAAGGTACACTTCAAATACCCTCGCCCTCCCAA 829
QY	176 -----GluGlyHisProSerGlnValGlySer----- 184
Db	830 CCCCAGCCCCCAGCAATCTACAGAGCACTGTCCGCATGAGCAGCTCTGACAGCAGTGA 889
QY	185 -----LysPheLeuLysAspLeuLysAsnLeuIleGluGlnLalaserasnSer 200
Db	890 AATTCGATGCTGATGTTTGACAGATGATGTTTTTACCTCGGATT---AAGCCAAATATACAT 946
QY	201 AsnAsnGlyLysProValIleLeuLeuSerHisSerLeuGlyIleuPheValLeuGln 220
Db	947 AATACGTCTCCAGATGATGCCAATAATGAACAATCTTACTAGATCAATTTGGATATCAT 1006
QY	221 LeuLeuAsnArgAsnProProSerTrpArgLysLysPheIleLysHisPheIleAlaLeu 240
Db	1007 GTTTCACAAC-----TTTTTGGCTGTG 1027

```
QY 241 SerAlaProTPrGlyGlyAlaIleAspGluMetTyrThrPheAlaSerGlyAsnThrLeu 260
    |||
    :|:|:|
    :|:|:|
Db 1028 AGCAATAGATATATGAAACCGAGACCTAAAGTATCTGTATAGATTAAGCACATAGCTTG 1087
    |||
QY 261 GlyValProLeuValAspProLeuValArgAspGluGlnArgSerSerGluSerAsn 280
    |||
    :|:|:|
    :|:|:|
Db 1088 GGTTTA----- 1093
QY 281 LeuTrpLeuLeuProAsnProLysIlePheGlyProGlnLysProIleValIleThrPro 300
    |||
    :|:|:|
    :|:|:|
Db 1094 -----CAGCTTCGTGGATGTA 1111
QY 301 IleArgProTyrSerAlaHisAspMetValAspPheLeuLysAspIleGlyPheProGlu 320
    :|:|
    :|:|:|
    :|:|:|
Db 1112 GTTCACAGTCATCAGACATATATGTCACATGAGCTCAAT-----GGCTTT----- 1159
QY 321 GlyValTyrProTyrGluThrArgIleLeuProLeuIleGlyAsnIleLysAlaProGln 340
    |||
    :|:|:|
    :|:|:|
Db 1159 ----- 1159
QY 341 ValProIleThrCysIleMetGlyThrGlyValGlyThrLeuGlnThrLeuPheTyrGly 360
    |||
    :|:|:|
    :|:|:|
Db 1160 -----GATATTGGCCAGAGTTTCACAGATCTTCTTCAATGCT 1198
QY 361 LysGlyAsp-----PheAspGluArgProGluIleSerTyrGlyAsp 374
    |||
    :|:|:|
    :|:|:|
Db 1199 ---GGAGACCGAGAGGATACCATAGTTGTGGATAGCAGG---CTGTTCAACTATGCCAAT 1252
QY 375 GlyAspGlyThrValAsnLeuValSerLeuLeuAlaLeuGlnSerLeuTrpLysGlu 394
    :|:|
    :|:|:|
    :|:|:|
Db 1253 TGGGAG-----GTCTCTGCTTCCCTTCCACTTGAGGTGGCTAGAAAGAG 1303
QY 395 LysAsnGlnTyrLeuLysValValLysIleAspGlyValSerHisThrSerIleLeuLys 414
    |||
    :|:|:|
    :|:|:|
Db 1304 TATTAAC-----TTTGACGATTTTCGATTTGATGCAATACTTCTATGCTATGTTCAAT 1357
QY 415 AspGluValAlaLeu-----AsnGluIleValGlyGluIleThrSer 428
    :|:|
    :|:|:|
    :|:|:|
Db 1358 CATGCAATCAATATGGATTTTACAGGAACATAATAGATATTTTCAGCGGCTACACAGAT 1417
QY 429 IleAsnSerHisAlaGluLeuGluLysSerAsn 439
    :|:|:|
    :|:|:|
    :|:|:|
Db 1418 GTTGATGCTGTGCTATTTAATGTTGGCCAAT 1450
```

Search completed: February 15, 2003, 15:22:53
Job time : 77 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 15, 2003, 14:56:22 ; Search time 65 Seconds

(without alignments)
3471.237 Million cell updates/sec

Title: US-09-857-612A-14

Perfect score: 2333

Sequence: 1 MKKEQEGELKIEVALTVTV.....GETTSINSHALGLSNLFGC 443

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 424239 segs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+.p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09857612/rnatc_11022003_111820_29961/app_query.fasta_1.583
-DB=Published_Applications_NA -QEMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09857612.qcgn_1_1_28.qrunat_11022003_111820_29961
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published_Applications_NA:*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12:	/cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1367	58.6	9	US-09-938-842A-376
2	610	26.1	9	US-09-978-295A-156
3	610	26.1	9	US-09-978-697-156
4	610	26.1	9	US-09-978-192A-156

5	610	26.1	2680	9	US-09-999-832A-156	Sequence 156, App
6	610	26.1	2680	9	US-09-978-189-156	Sequence 156, App
7	541	23.2	1744	10	US-09-919-497-30	Sequence 30, Appl
8	541	23.2	1744	10	US-09-880-107-2245	Sequence 2245, Ap
9	433	18.6	1048	9	US-09-998-059-8	Sequence 8, Appl1
10	433	18.6	1059	9	US-09-998-059-20	Sequence 20, Appl
11	433	18.6	1080	9	US-09-998-059-27	Sequence 27, Appl
12	411.5	17.6	2180	9	US-10-001-054-7	Sequence 7, Appl1
13	169.5	7.3	248	10	US-09-923-876-2515	Sequence 2515, Ap
14	151.5	6.5	527	10	US-09-867-701-5139	Sequence 5139, Ap
15	142	6.1	1482	9	US-09-938-842A-1935	Sequence 1935, Ap
16	134	5.7	266	10	US-09-960-352-2549	Sequence 2549, Ap
17	112	4.8	900	9	US-09-738-626-94	Sequence 94, Appl
18	110.5	4.7	393	10	US-09-960-352-2177	Sequence 94, Appl
19	104	4.5	954	9	US-09-938-842A-2479	Sequence 2479, Ap
20	103.5	4.4	1014	9	US-09-738-626-2470	Sequence 2470, Ap
21	101	4.3	2305	10	US-09-896-578-1	Sequence 1, Appl1
22	100.5	4.3	1029	10	US-09-896-578-3	Sequence 3, Appl1
23	100.5	4.3	1533	10	US-09-729-402-23	Sequence 23, Appl
24	100	4.3	20633	10	US-09-070-927A-276	Sequence 276, Appl
25	97.5	4.2	1041	9	US-10-027-805-24	Sequence 24, Appl
26	97.5	4.2	1041	10	US-09-903-410-24	Sequence 24, Appl
27	97.5	4.2	3042	9	US-09-738-626-1156	Sequence 1156, Ap
28	97.5	4.2	249487	9	US-10-026-188-3	Sequence 3, Appl1
29	97	4.2	1014	9	US-09-738-626-2192	Sequence 2192, Ap
30	96.5	4.1	1854	9	US-09-738-626-2266	Sequence 2266, Ap
31	96	4.1	9510	9	US-10-114-170-256	Sequence 256, Ap
32	95.5	4.1	1868	9	US-09-981-353-18	Sequence 18, Appl
33	95	4.1	263	10	US-09-878-574-6764	Sequence 6764, Ap
34	94	4.0	2416	10	US-09-748-739A-5	Sequence 5, Appl1
35	93.5	4.0	235	10	US-09-960-352-2308	Sequence 2308, Ap
36	93	4.0	3045	8	US-08-910-386A-16	Sequence 16, Appl
37	92.5	4.0	957	9	US-09-938-842A-316	Sequence 316, Appl
38	92	3.9	2191	10	US-09-729-402-22	Sequence 22, Appl
39	92	3.9	6544	10	US-09-764-847-1544	Sequence 1544, Ap
40	91.5	3.9	1170	10	US-09-815-242-6314	Sequence 6314, Ap
41	91.5	3.9	3029	9	US-09-834-975-811	Sequence 811, Appl
42	91	3.9	1239	9	US-09-938-842A-2482	Sequence 2482, Ap
43	91	3.9	1344	9	US-09-938-842A-2137	Sequence 2137, Ap
44	90	3.9	2381	10	US-09-880-107-2271	Sequence 2271, Ap
45	90	3.9	2416	10	US-09-748-739A-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-09-938-842A-376
: Sequence 376, Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Kreps, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: SAME-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: TITLE OF INVENTION: SAME-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: FILE REFERENCE: SCRIPT300-3
: CURRENT APPLICATION NUMBER: US/09/938, 842A
: PRIOR FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 60/227, 866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264, 647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300, 111
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 376
: LENGTH: 1242
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
: US-09-938-842A-376
Alignment Scores:

Pred. No.: 7,56e-155 Length: 1242
 Score: 1367.00 Matches: 28
 Percent Similarity: 77.26% Conservative: 58
 Best Local Similarity: 63.08% Mismatches: 83
 Query Match: 58.59% Indels: 10
 DB: 9 Gaps: 4

US-09-857-612a-14 (1-443) x US-09-938-842a-376 (1-1242)

QY LeuSerLeuLeuCythrCysGlyAlaSerAsnLeuAspProLeuIleLeuIleProGly 44
 DB 1 ARGACCTGATGTGTCACCTGTGGGTAGCAACGTTACCTTTGATTCTGGTCCAGCA 60
 QY 45 AsnGlyGlyAsnGlnLeuGluAlaArgLeuThrAsnGlnTyrLysProSerThrPheIle 64
 DB 61 AACGGAGTACACGCTAGACGCTGACAGCAAGATACAAAGCAAGTACTGCTCG 120
 QY 65 CysGluSerTrp---TyrProLeuIleLysLysAsnGlyTrpPheArgLeuTrpPhe 83
 DB 121 TGTACGAGCTGGTATATCCGATTCTATAGAGAGAGTGCTGGATGTTTAGCGTATGCTTC 180
 QY 84 AspSerSerValIleLeuAlaProPheThrGlnCysPheAlaGluArgMetThrLeuHis 103
 DB 181 GATCGACGAGCTGTATTGTCTCCCTCACCAGGTGCTTCAGCGATGATGATGTTGAC 240
 QY 104 TyrHisGlnGluLeuAspAspTyrPheAsnThrProGlyValGluThrArgValProHis 123
 DB 241 TATGACCCCTGATTTGATGATTACCAAAATGCTCGCTGTCACAAACCGGCTCTCAT 300
 QY 124 PheGlySerThrAsnSerLeuLeuTyrLeuAsnProArgLeuLysHisIleThrGlyTyr 143
 DB 301 TTGGGTTCAACAAATCCTATCTATACCTCGACCCCTCGTCCGAGATGCCACATCTTAC 360
 QY 144 MetAlaProLeuValAspSerLeuGln---LysLeuGlyTyrAlaAspGlyGluThrLeu 162
 DB 361 ATGGAAACATTTGGTGAACCTCTAGAGAAAATAAGCGGTATCTTACAGCAACCAACATC 420
 QY 163 PheGlyAlaProTyrAspPheArgTyrGlyLeuAlaAlaGlyGlyHisProSerGlnVal 182
 DB 421 CTAGGAGCTCCCATATGATTTTCAGGTACGGCTGGCTGCTTCGGCCACCCGCTGTGA 480
 QY 183 GlySerLysPheLeuLysAspLeuLysAsnLeuIleGluGluAlaSerAsnSerAsnAsn 202
 DB 481 GCCTCACATGCTCTACAAAGACCTCAAAACATGGTGGAAAAAATAGACGAGCAACGAA 540
 QY 203 GlyLysProValIleLeuLeuSerHisSerLeuGlyGlyLeuPheValLeuGlnLeu 222
 DB 541 GGAAGGCCAGTACTACTCTCTCCATAGCTAGAGAGACTTTTCTCTCATTTCTCTC 600
 QY 223 AsnArgAsnProProSerTrpArgLysLysPheIleLysHisPheIleAlaLeuSerAla 242
 DB 601 AACCTACACACCCCTTACATGAGCGCCGCAAGTACATCAACACATTTGTCACATCGCTGC 660
 QY 243 ProTrpGlyGlyAlaIleAspLeuLysLysPheIleLysHisPheIleAlaLeuSerAla 262
 DB 661 CCATGGGGGGGAGCATCTCTCAGATGAGACATTTGCTTGGCAACACATCGCTGTC 720
 QY 263 ProLeuValAspProLeuLeuValArgAspGluGlnArgSerSerGlnSerAsnLeuTrp 282
 DB 721 CCTTATAGTTAACCTTTGCTGGTTCAGACGCGCATCAGAGACCTCCGAGGTAACCAATGG 780
 QY 283 LeuLeuProAsnProLysIlePheGlyProGln---LysProIleValIleThrProIle 301
 DB 781 CTACTTCATCTACCAAAAGTGTTCACAGCAGACACTAAACCCCTTGCTACTCCCGAG 840
 QY 302 ArgProTyrSerAlaHisAspMetValAspPheLeuLysAspIleGlyPheProGluGly 321
 DB 841 GTTAACTACACAGCTTAGAGATGATGCGTTTTCGACAGACTTGGAATTCACAAAGCA 900
 QY 322 ValTyrProTyrGluThrArgIleLeuProLeuIleGlyAsnIleLysAlaProGlnVal 341
 DB 901 GTTGTGCTTACAGACAGAGAGTGTGCTTTTACAGAGAGACACTGATGACTCGGAGAGTG 960

QY 342 ProIleThrCysIleMetGlyThrGlyValGlyThrLeuGluThrLeuPheTyrGlyLys 361
 DB 961 CCAGTCACTTGGCATATATGGGAGAGGATGATACACCCGAGGTTTGTATGAGAAAA 1020
 QY 362 GlyAspPheAspGluArgProGluIleSerTyrGlyAspGlyAspGlyThrValAsnLeu 381
 DB 1021 GGAGATTTGGATGACACACAGAGATTAAGTATGAGATGAGATGGACGCTTAATTGG 1080
 QY 382 ValSerLeuLeuAlaLeuGlnSerLeuTyrLysGluGluLysAsnGlnTyrLeuLysVal 401
 DB 1081 GCGAGCTTAGACAGCTTGTG-----AAAGTCATAGCTTGAACACC 1119
 QY 402 ValLysIleAspGlyValSerHisThrSerIleLeuLysAspGluValAlaLeuAsnGlu 421
 DB 1120 GTAGAGATTTGATGAGCTTTCGATACATCTAATTAAGACAGAGATTCGACTTAAGAG 1179
 QY 422 IleValGlyGluIleThrSerIleAsn 430
 DB 1180 ATTATGACGAGATTTCATTTATTAT 1206

RESULT 2
 US-09-978-295A-156
 Sequence 156, Application US/09978295A
 Patent No. US20020156006A1
 GENERAL INFORMATION:
 APPLICANT: Ashkenazi, Avi
 APPLICANT: Baker Kevin P.
 APPLICANT: Botstein, David
 APPLICANT: Desnoyers, Luc
 APPLICANT: Eaton, Dan
 APPLICANT: Ferrara, Napoleon
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerber, Hanspeter
 APPLICANT: Gottard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, J. Christopher
 APPLICANT: Gurney, Austin L.
 APPLICANT: Hillan, Kenneth J.
 APPLICANT: Kljavin, Ivar J.
 APPLICANT: Kuo, Sophia S.
 APPLICANT: Napier, Mary A.
 APPLICANT: Pan, James:
 APPLICANT: Paonli, Nicholas F.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Shelton, David L.
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William I.
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 FILE REFERENCE: P2630PIC11
 CURRENT FILING DATE: 2001-10-15
 PRIOR APPLICATION NUMBER: 09/918585
 PRIOR FILING DATE: 2001-07-30
 PRIOR APPLICATION NUMBER: 60/062250
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/064249
 PRIOR FILING DATE: 1997-11-03
 PRIOR APPLICATION NUMBER: 60/065311
 PRIOR FILING DATE: 1997-11-13
 PRIOR APPLICATION NUMBER: 60/066364
 PRIOR FILING DATE: 1997-11-21
 PRIOR APPLICATION NUMBER: 60/077450
 PRIOR FILING DATE: 1998-03-10
 PRIOR APPLICATION NUMBER: 60/077632
 PRIOR FILING DATE: 1998-03-11
 PRIOR APPLICATION NUMBER: 60/077641
 PRIOR FILING DATE: 1998-03-11

Page 5

[illegible]

PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083356
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083455
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083454
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083555
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084411
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085399
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15

TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630P1C7
CURRENT APPLICATION NUMBER: US/09/978,189
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071

PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07

;; PRIOR APPLICATION NUMBER: 60/108849
;; PRIOR FILING DATE: 1998-11-18
;; PRIOR APPLICATION NUMBER: 60/112420
;; PRIOR FILING DATE: 1998-12-15
;; PRIOR APPLICATION NUMBER: 60/113296
;; PRIOR FILING DATE: 1998-12-22
;; PRIOR APPLICATION NUMBER: 60/115554
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115558
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/116533
;; PRIOR FILING DATE: 1999-01-20
;; PRIOR APPLICATION NUMBER: 60/123618
;; PRIOR FILING DATE: 1999-03-10
;; PRIOR APPLICATION NUMBER: 60/131294
;; PRIOR FILING DATE: 1999-04-07
;; PRIOR APPLICATION NUMBER: 60/140650
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/141037
;; PRIOR FILING DATE: 1999-06-23
;; PRIOR APPLICATION NUMBER: 60/144758
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 60/162506
;; PRIOR FILING DATE: 1999-10-29
;; PRIOR APPLICATION NUMBER: 60/170262
;; PRIOR FILING DATE: 1999-12-09
;; PRIOR APPLICATION NUMBER: 60/187202
;; PRIOR FILING DATE: 2000-03-03
;; PRIOR APPLICATION NUMBER: 60/209832
;; PRIOR FILING DATE: 2000-06-05
;; PRIOR APPLICATION NUMBER: 60/222887
;; PRIOR FILING DATE: 2000-09-15
;; PRIOR APPLICATION NUMBER: 09/180997
;; PRIOR FILING DATE: 1998-11-19
;; PRIOR APPLICATION NUMBER: 09/218517
;; PRIOR FILING DATE: 1998-12-22
;; PRIOR APPLICATION NUMBER: 09/284291
;; PRIOR FILING DATE: 1999-04-12
;; PRIOR APPLICATION NUMBER: 09/380137
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/380138
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/380913
;; PRIOR FILING DATE: 1999-09-09
;; PRIOR APPLICATION NUMBER: 09/403297
;; PRIOR FILING DATE: 1999-10-18
;; PRIOR APPLICATION NUMBER: 09/423741
;; PRIOR FILING DATE: 1999-11-10
;; PRIOR APPLICATION NUMBER: 09/709238
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 09/802706
;; PRIOR FILING DATE: 2001-03-09
;; PRIOR APPLICATION NUMBER: 09/866034
;; PRIOR FILING DATE: 2001-05-25
;; PRIOR APPLICATION NUMBER: 09/872035
;; PRIOR FILING DATE: 2001-06-01
;; PRIOR APPLICATION NUMBER: 09/882636
;; PRIOR FILING DATE: 2001-06-14
;; PRIOR APPLICATION NUMBER: 09/918585
;; PRIOR FILING DATE: 2001-07-30
;; PRIOR APPLICATION NUMBER: 09/924419
;; PRIOR FILING DATE: 2001-08-06
;; PRIOR APPLICATION NUMBER: 09/927796
;; PRIOR FILING DATE: 2001-08-06
;; PRIOR APPLICATION NUMBER: 09/929404
;; PRIOR FILING DATE: 2001-08-13
;; PRIOR APPLICATION NUMBER: 09/941992
;; PRIOR FILING DATE: 2001-08-28
;; PRIOR APPLICATION NUMBER: 09/946374
;; PRIOR FILING DATE: 2001-09-04
;; PRIOR APPLICATION NUMBER: PCT/US98/18824
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: PCT/US99/00106

;; PRIOR FILING DATE: 1999-01-05
;; PRIOR APPLICATION NUMBER: PCT/US99/05028
;; PRIOR FILING DATE: 1999-03-08
;; PRIOR APPLICATION NUMBER: PCT/US99/08615
;; PRIOR FILING DATE: 1999-04-20
;; PRIOR APPLICATION NUMBER: PCT/US99/12252
;; PRIOR FILING DATE: 1999-06-02
;; PRIOR APPLICATION NUMBER: PCT/US99/20111
;; PRIOR FILING DATE: 1999-09-01
;; PRIOR APPLICATION NUMBER: PCT/US99/20594
;; PRIOR FILING DATE: 1999-09-08
;; PRIOR APPLICATION NUMBER: PCT/US99/28313
;; PRIOR FILING DATE: 1999-11-30
;; PRIOR APPLICATION NUMBER: PCT/US99/28551
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/28634
;; PRIOR FILING DATE: 1999-12-01
;; PRIOR APPLICATION NUMBER: PCT/US99/30095
;; PRIOR FILING DATE: 1999-12-16
;; PRIOR APPLICATION NUMBER: PCT/US99/30999
;; PRIOR FILING DATE: 1999-12-20
;; PRIOR APPLICATION NUMBER: PCT/US00/00376
;; PRIOR FILING DATE: 2000-01-06
;; PRIOR APPLICATION NUMBER: PCT/US00/03565
;; PRIOR FILING DATE: 2000-02-11
;; PRIOR APPLICATION NUMBER: PCT/US00/04341
;; PRIOR FILING DATE: 2000-02-18
;; PRIOR APPLICATION NUMBER: PCT/US00/04342
;; PRIOR FILING DATE: 2000-02-18
;; PRIOR APPLICATION NUMBER: PCT/US00/05841
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: PCT/US00/06684
;; PRIOR FILING DATE: 2000-03-15
;; PRIOR APPLICATION NUMBER: PCT/US00/08439
;; PRIOR FILING DATE: 2000-03-30
;; PRIOR APPLICATION NUMBER: PCT/US00/13705
;; PRIOR FILING DATE: 2000-05-17
;; PRIOR APPLICATION NUMBER: PCT/US00/14042
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: PCT/US00/14941
;; PRIOR FILING DATE: 2000-05-30
;; PRIOR APPLICATION NUMBER: PCT/US00/15264
;; PRIOR FILING DATE: 2000-06-02
;; PRIOR APPLICATION NUMBER: PCT/US00/22031
;; PRIOR FILING DATE: 2000-08-11
;; PRIOR APPLICATION NUMBER: PCT/US00/23522
;; PRIOR FILING DATE: 2000-08-23
;; PRIOR APPLICATION NUMBER: PCT/US00/30873
;; PRIOR FILING DATE: 2000-11-10
;; PRIOR APPLICATION NUMBER: PCT/US00/32678
;; PRIOR FILING DATE: 2000-12-01
;; PRIOR APPLICATION NUMBER: PCT/US01/06520
;; PRIOR FILING DATE: 2001-02-28
;; PRIOR APPLICATION NUMBER: PCT/US01/06666
;; PRIOR FILING DATE: 2001-03-01
;; PRIOR APPLICATION NUMBER: PCT/US01/17092
;; PRIOR FILING DATE: 2001-05-25
;; PRIOR APPLICATION NUMBER: PCT/US01/17800
;; PRIOR FILING DATE: 2001-06-01
;; PRIOR APPLICATION NUMBER: PCT/US01/19692
;; PRIOR FILING DATE: 2001-06-20
;; PRIOR APPLICATION NUMBER: PCT/US01/21066
;; PRIOR FILING DATE: 2001-06-29
;; PRIOR APPLICATION NUMBER: PCT/US01/21735
;; PRIOR FILING DATE: 2001-07-09
;; PRIOR APPLICATION NUMBER: PCT/US01/27099
;; PRIOR FILING DATE: 2001-08-29
;; NUMBER OF SEQ ID NOS: 91
;; SEQ ID NO 7
;; LENGTH: 2180
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-001-054-7

Alignment Scores:

Pred. No.:	2,17e-39	Length:	2180
Score:	411.50	Matches:	92
Percent Similarity:	56.08%	Conservative:	51
Best Local Similarity:	36.08%	Mismatches:	97
Query Match:	17.64%	Indels:	15
DB:	9	Gaps:	8

US-09-857-612A-14 (1-443) x US-10-001-054-7 (1-2180)

```

QY 167 TYRASPHEARGTYRIGLYLEUALAAGLUGLHISPROSEGLINVALGYSERLYSPHE 186
    |||:|||||
DB 1 TATGACTGGCGCGAGACC-----CCAAATGAAAACGGCCCTACTTC 42
QY 187 LEULYSASPLEULYSANLEULIEGLUGLUALASERASNSRASNGLYSPROVAL 206
    |||:|||||
DB 43 CTGGCC--CTCCGCGAGATGATCGAGAGATGACCTGATGAGGCGC--CCGTCG 96
QY 207 ILEULEUSERHISSEULEUGLIGLYLEUPHEVALLEUGLNULEUASNAARGASPRO 226
    |||:|||||
DB 97 GTGCTGTGTCGCCACAGTATGGGCAACATGTACACGCTTCTCTGCGAGCGGACCGC 156
QY 227 PROSETRPARGLYSPHEILEULYSASPHLEILEALALEUSERALAPROTPRGILY 246
    |||:|||||
DB 157 CAGGCTCGAAGAGACAATATATCGCGCTTGTCTCACTGGGTCGCGCCCTGGGCGG 216
QY 247 ALAILEASPLUMETRYTHRPHLEALASERGLYASNTHLEUGLYVALPROLEUVALASP 266
    |||:|||||
DB 217 GTGGCCAGAACCTCGCGCTCTCTGCTTCAAGAGACACACACCGCATCCGATCGGG 276
QY 267 PROLEULEVALARGASPLUGLUGLARGSERSEGLUSERASNULEUPROLEUVALASP 286
    |||:|||||
DB 277 CCCCTAAGATCGGACAGCAGCGGTCAGCTCTCCACGACGTCGCTGCGCTGCGCTAC 336
QY 287 PROLYSILEPHGLYPROGLINUSPROILEVALIETHRPROILEARGPROTYRSEALA 306
    |||:|||||
DB 337 AACTACACATGTCACCTGAGAGAGTTCGTCGACAGACACCCAAATGACATACACCTG 396
QY 307 HISASPMETVALASPHLEULYSASPLIEGLYPHEPROGLINLY--VALTYRPROTYR 325
    |||:|||||
DB 397 CGGACTACCGCAAGTCTTCTCCAGACATCGCTTGAAAGTGGCTCATGCGGCAG 456
QY 326 GLUTHRARGILEUPROLEULEGLYASNILELYSALAPROGLINVALPROILETHRYS 345
    |||:|||||
DB 457 GACACAGAA-----GGCTGTGTGAAGCCACACATGCCACTGCGTGCACGCTGCAC 510
QY 346 ILEMETGLYTHRGLYVALGLYTHLEUGLUTHRLEUPHETRGLYLYSGLYASPPHEASP 365
    |||:|||||
DB 511 CTTATGATGACTGGCTGCCACACACACACTCTCTACTAT-----GAGAGCTTCCCT 564
QY 366 GLUARG---PROGLULESERTYRGLYASPLGYSPLGYTHRYVALASNULEUVALSERLEU 384
    |||:|||||
DB 565 GACCGAGACCTTAATCTGCTTGTGTGACGGGATGACTGTGAACATGGAAGTGGC 624
QY 385 LEUALALEUGLINSERLEUPRYSGLUGLULYASNGLINTRYLEUYSVALVALYSTILE 404
    |||:|||||
DB 625 CTGCAATGCCAGGCC---TGGCAGAGCCGCCAGAGACCAACTGTTCGCGAGAGCTG 661
QY 405 ASPLGVALSERHISHTHRSERILELEULYSASPLUVALALALEU 419
    |||:|||||
DB 682 CCAGCGACGAGACATGAGATGTCGCCACACGACACACCTG 726

```

RESULT 13

```

US-09-923-876-2515
; Sequence 2515, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Ialoudi, Raghuath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON

```

```

; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 2515
; LENGTH: 248
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700160724H1
US-09-923-876-2515

```

Alignment Scores:

Pred. No.:	9,26e-12	Length:	248
Score:	169.50	Matches:	33
Percent Similarity:	67.16%	Conservative:	12
Best Local Similarity:	49.25%	Mismatches:	21
Query Match:	7.27%	Indels:	1
DB:	10	Gaps:	1

US-09-857-612A-14 (1-443) x US-09-923-876-2515 (1-248)

```

QY 150 SERLEUGLINSLEUGLYTRYALIASPLGLUTHRLEUPHEGLYALAPROTYRASPHE 169
    |||:|||||
DB 51 GCACGTGAGACATGGATGATCCGAGACGAGACACCATGTCGAGCGCCCTACGACTTC 110
QY 170 ARGTYRGLYLEUALAAGLUGLHISPROSEGLINVALGYSERLYSPHEULEYASP 189
    |||:|||||
DB 111 CGCTACCGCGCGCGCTCCCGGCGCAACGTCGAGGTGTACTCCGCTCAAGAG 170
QY 190 LEULYSANLEULIEGLUGLUALASERASNSRASNGLYSPROVALIILEULEU 209
    |||:|||||
DB 171 CTGATGAGAGCTGTCGAGCGCGCAAGCGAGAGACCGGAAGAGCGGCTGATC--CTC 227
QY 210 SERHISERLEUGLYGLYLEU 216
    |||:|||||
DB 228 GGCACACAGCTTCGCGCGCATG 248

```

RESULT 14

```

US-09-867-701-5139
; Sequence 5139, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Agilate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121,497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5139
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(527)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-5139

```

Alignment Scores:

Pred. No.:	4,56e-09	Length:	527
Score:	151.50	Matches:	44
Percent Similarity:	51.19%	Conservative:	42
Best Local Similarity:	26.19%	Mismatches:	77

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2003, 14:02:12 ; Search time 1680 seconds
(without alignments)
4270.596 Million cell updates/sec

Title: US-09-857-612a-14
Perfect score: 2333
Sequence: 1 MKKEQEGSLKIEVATLVTV.....GEITSINSMAELGLSNLFSG 443

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model -DEV=xlh
-O=/cgn2.1/USPTO/spool/US09857612/runat_11022003_111818_29838/app_query.fasta_1.583
-DB=EST -OFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pt0 -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09857612.ecgn_1_1.899_etrunat_11022003_111818_29838 -NCPU=6 -ICPU=3
-NO_XLUPX -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST :
1: em_estba :
2: em_estlum :
3: em_estlin :
4: em_estlmu :
5: em_estcov :
6: em_estpl :
7: em_estro :
8: em_hlc :
9: gb_est1 :
10: gb_est2 :
11: gb_hlc :
12: gb_est3 :
13: gb_est4 :
14: gb_est5 :
15: em_estfun :
16: em_estom :
17: gb_gss :
18: em_gss_hum :
19: em_gss_inv :
20: em_gss_pln :
21: em_gss_vrt :
22: em_gss_fun :
23: em_gss_mam :
24: em_gss_mus :
25: em_gss_other :
26: em_gss_pro :
27: em_gss_rtd :

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	904	38.7	537	10	AW759200
2	903	38.7	532	13	BI973838
3	868	37.2	650	12	BG524903
4	809	34.7	648	14	BQ852109
5	784	33.6	630	14	BQ855578
6	774	33.2	507	14	BM885973
7	744	31.9	673	10	BE322181
8	694	29.7	587	12	BG523648
9	650	27.9	618	10	AV441327
10	624	26.7	635	12	BG522086
11	614	26.3	596	12	BG525847
12	590.5	25.3	724	12	BG127829
13	573	24.6	616	12	BG526525
14	561	24.0	537	10	BE355801
15	560.5	24.0	527	14	BQ864610
16	543	23.3	593	17	BH777228
17	518	22.2	1350	11	AK004914
18	510.5	21.9	529	10	AV442635
19	509.5	21.8	600	10	AV827187
20	500	21.4	657	13	BJ256228
21	485	20.8	1141	13	BM453791
22	444.5	19.1	977	14	BO711543
23	426	18.3	673	13	BJ261764
24	424.5	18.2	731	13	BI406095
25	420.5	18.0	682	14	BM770153
26	420.5	18.0	737	14	BM770144
27	420.5	18.0	594	9	AL524311
28	413	17.7	910	10	AV528103
29	412.5	17.7	993	12	BG829912
30	410.5	17.6	676	17	BH578169
31	405	17.4	608	14	BM771194
32	405	17.4	662	14	BM771402
33	400	17.1	928	14	BQ707084
34	397.5	16.8	767	11	AV112503
35	391	16.8	401	13	BI139236
36	386	16.5	582	14	BM771795
37	385	16.5	637	14	BM459510
38	378.5	16.2	827	9	AI047852
39	376.5	16.1	944	13	BI330091
40	373.5	16.0	419	10	AV816862
C	41	372.5	837	13	BI327594
42	372	15.9	841	12	BE310537
43	369	15.8	551	14	BM770078
44	362	15.5	570	14	BM771852
45	355	15.2	370	10	BE355900

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	AW759200	527 bp	mrna	linear	EST 03-DEC-2001
AW759200	LOCUS	AW759200	527 bp	mrna	linear	EST 03-DEC-2001	
AW759200	DEFINITION	Gm-cl027-3587 5' similar to TR:092M42 Q92M42 F17L21.28. ;, mRNA					
AW759200	ACCESSION	AW759200					
AW759200	VERSION	AW759200.1	GI:7691066				
AW759200	KEYWORDS	EST.					
AW759200	SOURCE	soybean.					
AW759200	ORGANISM	Glycine max					
AW759200		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
AW759200		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;					
AW759200		Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;					
AW759200		Glycine.					

REFERENCE

1 (bases 1 to 527)

AUTHORS

Shoemaker, R., Keim, P., Vodka, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE

Public Soybean EST Project

JOURNAL

Unpublished (1999)

COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1800
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 for further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 900 Std Error: 0.00
High quality sequence stop: 435.

FEATURES

source

1..527
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl027-3587"
/clone_1db="Gm-cl027"
/tissue_type="cotyledons of 3- and 7-day-old Williams
seedlings"

/lab_host="DH10B"
/note="Vector: pBluescript II SK(+). Site 1: EcoRI; Site 2: XhoI. This cDNA library was constructed from mRNA isolated from cotyledons of 3- and 7-day-old Williams seedlings which were propagated on paper towels with distilled water. The cotyledons were flash-frozen in liquid nitrogen, then lyophilized for 72 hours. Unequal amounts of mRNA was used for cDNA synthesis. Stratagene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An anchor nucleotide (V-A, C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGACTCTCTGAG(T)18] to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GIBCOBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript(II) SK(+) digested with EcoRI and XhoI, and phosphorylated by Stratagene. 97% of the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=30). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

BASE COUNT 156 a 107 c 116 g 148 t
ORIGIN

Alignment Scores:

Pred. No.: 4,91e-102 Length: 527
Score: 904.00 Matches: 170
Percent Similarity: 98.29% Conservative: 2
Best Local Similarity: 97.14% Mismatches: 3
Query Match: 38.75% Indels: 0
DB: 10 Gaps: 0

US-09-857-612a-14 (1-443) x AW759200 (1-527)

QY 210 SerHisSerLeuGlyGlyLeuPheValLeuGlnLeuLeuAsnArgAsnProProSerThr 229
|||||
Db 3 TCCACACTTTAGAGGGCGCTATTGTCTTACACTACTAATAATAGAACCCCGCTTGG 62
QY 230 ArgLysLysPheIleLysHisPheIleAlaLeuSerAlaProTTPGlyGlyAlaIleAsp 249
Db 63 CGCAAAAATTCATCAACCACTTATGCTCTTTCAGCTCATGGGGGGGCTATAGAC 122
QY 250 GlnMetTyrThrPheAlaSerGlyAsnThrLeuGlyValProLeuValAspProLeuLeu 269
Db 123 GAATGTACACTTTGCTGCTTGGCAACACTTTGGAGTGGCCCTAGTGGACCTTTATTA 182
QY 270 ValArgAspGlnArgSerSerGluSerAsnLeuThrLeuLeuProAsnProLysIle 289
Db 183 GTGAGGATGACAAAGAAAGCTCCGAGATACCTTTGGCTTTGGCTTAACCAAAATTT 242
QY 290 PheGlyProGlnLysProIleValIleThrProIleArgProTyrSerAlaHisAspMet 309
Db 243 TTTGGTCTCAAAAACCAATAGTATTAATCAATAATAGACCTATTCAGCTCATGACATG 302
QY 310 ValAspPheLeuLysAspIleGlyPheProGlnGlyValTyrProTyrGluThrArgIle 329
Db 303 GTTGATTTTCTTAAAAAGACATTGTTTCTTCAAGGGCTTTATCTTATGAACAGCAATT 362
QY 330 LeuProLeuIleGlyAsnIleLysAlaProGlnValProIleThrCysIleMetGlyThr 349
Db 363 CTACCTCTGATAGGAGACATTAAGACACACACAGCTGCTTATTAATGATATGAGGAACG 422
QY 350 GlyValGlyThrLeuGluThrLeuPheTyrGlyLysGlyAspPheAspGluArgProGlu 369
Db 423 GGAATGGGAACCTTGGAACCACTTTTATGGAAGGTGATTTTGATGAAGCGGCAGAA 482
QY 370 IleSerTyrGlyAspGlyAspGlyThrValAsnLeuValSerLeu 384
Db 483 ATATCATATGGGATGCGATGATGACGCTGATCTTGAGACTTG 527

RESULT 2 532 bp mRNA linear EST 30-NOV-2001
BI973838
LOCUS sa193608.y1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE
DEFINITION ID: Gm-cl065-8607 5' similar to TR:092M42 Q92M42 F1L21.28. ; mRNA
sequence.

ACCESSION BI973838
VERSION BI973838.1 GI:16348243
KEYWORDS EST.

SOURCE

soybean.
Glycine max

REFERENCE

1 (bases 1 to 532)

AUTHORS

Shoemaker, R., Keim, P., Vodka, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE

Public Soybean EST Project

JOURNAL

Unpublished (1999)

COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1800
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 for further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 421.
Location/Qualifiers

FEATURES

source

1. .532
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl065-8607"
/clone_lib="Gm-cl065"
/tissue_type="germinating shoots"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site.1: EcoRI; Site.2: XhoI. The cDNA library was constructed from mRNA isolated from germinating shoots of the cultivar Williams. The seeds were allowed to germinate for 24 hours prior to being cold stressed for 2 days at 4C. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 168 a 87 c 124 g 152 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 6,65e-102 Length: 532
Score: 903.00 Matches: 174
Percent Similarity: 98.87% Conservative: 1
Best Local Similarity: 98.31% Mismatches: 2
Query Match: 38.71% Indels: 0
DB: 13 Gaps: 0

US-09-857-612A-14 (1-443) x B1973838 (1-532)

QY 254 PheAlaSerGlyAsnThrLeuGlyValProLeuValAspProLeuLeuValArgAspGlu 273
|||||
DB 2 TTTGCATCTGCGCAACCTTGGAGAGCCCTGTGACCCCTTATTAGTAGAGGATGAA 61
|||||

QY 274 GlnArgSerSerGluSerAsnLeuTrpLeuLeuProAsnProValIlePheGlyProGln 293
|||||
DB 62 CAAAGAGCTCCGAGAGTAACTTGGCTTTGCTTACCCCAAAATTTTGGTCTCCAA 121
|||||

QY 294 LysProIleValIleThrProIleArgProTyrSerAlaHisAspMetValAspPheLeu 313
|||||
DB 122 AAACCAATAGTGAATCACTCAATAGACCTATTCAGCTCATGATGCTTGAATTTCTA 181
|||||

QY 314 LysAspIleGlyPheProGluGlyValTyrProTyrGluThrArgIleLeuProLeuIle 333
|||||
DB 182 AAGACATTTGGTTTCTGAGGGGTTTATCTTATGAAACAGAAATTCATCCCTGATA 241
|||||

QY 334 GlyAsnIleLysAlaProGlnValProIleThrCysIleMetGlyThrGlyValGlyThr 353
|||||
DB 242 GGGACATTAAGACACACAGTGCCTATTAAGTGAAGGAGGAGGGAGACC 301
|||||

QY 354 LeuGluThrLeuPheTyrGlyLysGlyAspPheAspGluArgProGluIleSerTyrGly 373
|||||
DB 302 TTGGAACAATGTTTATGGGAAGTGATTTGATGAACGGCAGAAATATNTATGGG 361
|||||

QY 374 AspGlyAspGlyThrValAsnLeuValSerLeuLeuAlaLeuGlnSerLeuTrpLysGlu 393
|||||
DB 362 GAGGTGATGAGGAACGGTGAACCTTGGTGGCCCTTCAATCACTATGAGAAAG 421
|||||

QY 394 GlnLysAspGlnTyrLeuLysValIleLysIleAspGlyValSerHisThrSerIleLeu 413
|||||
DB 422 GAGAAATAATCAATACCTTAAGTGGTTAAGATGATGGGTGCTCTACTTCAATACCT 481
|||||

QY 414 LysAspGlnValAlaLeuAsnGluIleValGlyGluIleThrSerIleAsn 430
|||||
DB 482 AAGGATGAAGTTGCACTAAATAGATAGTGAAGTATTCCTCAATATAT 532
|||||

RESULT 3
BG524903 650 bp mRNA linear EST 16-NOV-2001
LOCUS

DEFINITION 8-8 Stevia field grown leaf cDNA Stevia rebaudiana cDNA 5', mRNA sequence.
ACCESSION BG524903
VERSION BG524903.1 GI:16948350
KEYWORDS EST.
SOURCE Stevia rebaudiana.
ORGANISM Stevia rebaudiana
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Asteridae: euasterids II: Asterales: Asteraceae: Asteroideae; Helianthaceae; Stevia.
1 (bases 1 to 650)
REFERENCE Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P. Leaf ESTs from Stevia rebaudiana: A resource for gene discovery in diterpene synthesis
JOURNAL Unpublished (2001)
COMMENT Contact: Jim Brandle
Genomics and Biotechnology
Agriculture and Agri-Food Canada - SCPPRC
1391 Sandford St., London, Ontario, CANADA, N5V 4T3
Tel: 519 457 1470
Fax: 519 457 3997
Email: brandleje@am.agr.ca
Seq primer: T3 promoter primer.
Location/Qualifiers
1. .650
/organism="Stevia rebaudiana"
/strain="751/1501"
/cultivar="Landrace"
/db_xref="taxon:55670"
/clone_lib="Stevia field grown leaf cDNA"
/tissue_type="leaf"
/dev_stage="field grown, mid-size"
/lab_host="E. coli strain XL0LR"
/note="Vector: pBK-CMV; Site.1: EcoRI; Site.2: XhoI. This cDNA library was constructed from polyA+ enriched mRNA from field grown leaves. Mid-size actively growing leaves were collected and pooled from several plants and frozen immediately after harvesting in liquid nitrogen. The cDNA was prepared using an XhoI-poly(dT) linker-primer. An EcoRI adapter was ligated to the blunt end cDNA and the products were digested with EcoRI and XhoI enabling directional cloning into the lambda ZAP Express vector. The library was amplified using the host strain XL1-Blue MRF. Mass excision of the library was performed to obtain pBK-CMV phagemid clones in the host strain XL0LR. Single pass DNA sequencing was performed using the T3 promoter primer: 5' ATTACCTCTACCTAAAGGGA 3'. This library was constructed by Alex Richman."

BASE COUNT 171 a 164 c 143 g 171 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 2.18e-97 Length: 650
Score: 868.00 Matches: 160
Percent Similarity: 85.58% Conservative: 24
Best Local Similarity: 74.42% Mismatches: 31
Query Match: 37.21% Indels: 0
DB: 12 Gaps: 0

US-09-857-612A-14 (1-443) x BG524903 (1-650)

QY 79 PheArgLeuTrpPheAspSerSerValIleLeuAlaProPheThrGlnCysPheAlaGlu 98
|||||
DB 1 TTCAAGCTATGAGTTGCGAGGCGGCGTGTGTGGACCGGTGACCGGATGTTGGCCGAT 60
|||||

QY 99 ArgMetThrLeuHisTyrHisGlnGluLeuAspAspTyrPheAsnThrProGlyValGlu 118
|||||
DB 61 CGTATGACGCTTACTATGACCGAGATGCGATGATACAAAATAGTCCCGGAGTGGAG 120
|||||

QY 119 ThrArgValProHisPheGlySerThrAsnSerLeuLeuTyrLeuAsnProArgLeuLys 138
|||||
DB 121 ACTAGAGTGTCTGAGTTGGTGTCTACTCAATCGCTTCTACCTTGATGATCTTCTCAAG 180
|||||

Oy		139	HslieHhGlylYmeLxlarProleuValSerLeuA16ylYurAlAsP	158
Oy		139	HslieHhGlylYmeLxlarProleuValSerLeuA16ylYurAlAsP	158
Dd		181	CATATACAACTACTAGTGAACCATTTGGGGAATGTATACCAACCTTGCGTCACAAAGT	240
Oy		159	GlyGUtHrHeuPhedGLyALabProTyArSphEarGTYGLYLeuaLaalagLUgLnIs	178
Dd		241	ACCAGAAATCTTTTGGAGCCTCATATATGATTTCGTACGGGTAGCATCCGAAGACAT	300
Oy		179	ProSeGrInValGlySerTLysPheLeuYsaSRpleuLYAsnLeuILeGIuLuAlaser	198
Dd		301	CTTGTGATGTGGGCTCAACTTAAGCTCCAAAACSTPAACCACTTAATGCMAAACCAAC	360
Oy		199	AnSerASAnSGLyLSySProVALlleLeuLeuSerHisSerLeuGLYglyLeuPheval	218
Dd		361	AACACAATAAGGGGAGACTCCGGGTGATCTCCGCTGCCACAGTCTTGAGAGCCTGTCCGA	420
Oy		219	LeuGLInLeuLeuAsnArgAsnPProSeGrTParGLYsALysPheILelshIsPheIle	238
Dd		421	CTCCAAGTCCTCAACCGTTAACCCACCCTCATGGCCCAACATTAATTAACATTTTATAC	480
Oy		239	AlAleuSerAlaProTrpGLYglYALaleaRGlumeTYTThPheLasergLYasn	258
Dd		481	GGGTGTCCGCACCATGGGGGTGAACGGTTCAGAGATGTAACTTTTCTCTGGAGAC	540
Oy		259	ThrLeuGLYvalProLeuValAsPProLeuLeuValArgASPGLInArgSerSerglu	278
Dd		541	ACTCTGGGAGTTCACACTGCTGTGAACCCGTTGCTGTGCAGAAACGACACGAGTCTGAA	600
Oy		279	SerAsnLeuTrPLeuLeuProAsnProLYsilePheGLYProGLn	293
Dd		601	AGCAATTTATGGCTCATGTCACAGCAAACAAATTTNCTACACAG	645
RESULT 4				
B0852109				
LOCUS				
DEFINITION		B0852109	648 bp mRNA linear EST 14-AUG-2002	
ACCESSION		OGB17C13	.yg.ab1 QG_ABCDI lettuce salinas Lactuca sativa cDNA clone	
VERSION		B0852109		
KEYWORDS		B0852109.1	GI:22237574	
SOURCE			EST.	
ORGANISM			Lactuca sativa.	
REFERENCE			Lactuca sativa	
AUTHORS			Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
JOURNAL			Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;	
COMMENT			Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;	
TITLE			Lactuca.	
JOURNAL			1 (bases 1 to 648)	
COMMENT			Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Riesberg,L.,	
			Liu,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison	
			,P., Kolman,J.J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,	
			Church,S., Jackson,L. and Bradford,K.	
			Lettuce and Sunflower ESTs from the Compositae Genome Project	
			http://compenomics.ucdavis.edu/	
			Unpublished (2002)	
			Contact: Alexander Kozik [R.W.Michelmore]	
			Department of Vegetable Crops, R.W.Michelmore Lab	
			University of California at Davis (UCD)	
			Asmundson Hall, UCD, Davis, CA 95616, USA	
			Tel.: 1-(530)-742-1742	
			Fax: 1-(530)-752-9659	
			Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]	
			belongs to contig QG_Coln19973, see http://cgpbh.ucdavis.edu/	
			for details.	
FEATURES			Plate: OGB17 row: C column: 13.	
SOURCE			Location/Qualifiers	
			1..648	
			/organism="Lactuca sativa"	
			/cultivar="Salinas"	
			/db_xref="taxon:4236"	
			/clone="QGB17C13"	
			/lab_host="QG_ABCDI lettuce salinas"	
			/lab_host="E.coli"	

QY	177	GLYHISPROSERGLINVALGILSERLYSPHELEULYSASPHEULYSASNLEULIEGLUGLU	196
Db	1	GGGCATCCCTTGGAGCGTGGCTCCACCTTCCTCCCAAAACATTAACAGATTAAATCAGAGAA	60
QY	197	ALASERASERASASNGLYSPROYALILEULEUSERHISSEITREUGILYGLILEU	216
Db	61	GCAAGTGTATCAATAGTGGAAGAACCCGTCATCTTCATCTCCCATAGTCTCGTGCGCTC	120
QY	217	PHEVALLEUGLILEULLEUASNAARGANPROPPOSERTATRGLYSPHEILEULYSHIS	236
Db	121	TTTCGTCCTCCAACTCCTCAACCGCAACCGCTCTCGGCGGAACATATACATAAAGCAC	180
QY	237	PHLEIALALEUSERALAPROTIRPILYALALILEASPGILMETYRTYTHPHEALASER	256
Db	181	TTTCATCGCAGTACGCGGACCATGCGGTGGACGGTTGACAGAGATGTGACCTTTGCATTC	240
QY	257	GLYASNTHIRLEUGLYVALPROLEULVALSPROLEULVALARGASPLUGLNAIRGSE	276
Db	241	GGGAATTCACTAGAGAGTCCCTCGTAAACCCCTTGCTGTGTAGAAACGAACACGAGAC	300
QY	277	SERGIUSERASNLEUTRPLEULEUPROASNPOLYSILIEPHEGLYPROGLINLSPROILE	296
Db	301	TCAGAGACCAACTATATGCTATTCGACAGTCAACAAAGCTTCCCGCAGAGAGAAACCACTT	360
QY	297	VALIETHIRPROLEIARGPROTYISERALAHISASMETVALASPPHEULEULYSAPILE	316
Db	361	GTTGTTACCAAAACCTGACATTAATCTCTTGTGTATTTCAAGATTTTGAAGAATATAC	420
QY	317	GLYPHEPRGILUGLYVALTYTRPROTYRILUNHARGILIELEUPROLEULIEGLYSNILE	336
Db	421	GGCTTTCTCGTAGAGGGGTCCACCTTACCAAGAAACCGCTTGCCCTGTGTGAGAGATTG	480
QY	337	LYSALAPROGLINVALPROILETHRCYSILMETGLYTHRGLYVALGILYTHIRLEULUNH	356
Db	481	GCTGCACCAAGAAATTCGCTAATCATGTATTTATCGGAAGTGGAGTTCAACACCGCAATCT	540
QY	357	LEUPHEUTRGILYSGLYASPPHEASPLUARGPROLUILESERTYRGLYASPLIYASP	376
Db	541	TTTGATTTACGGAAGAAAGAGGTTTGATTAAGACAGCGGAGATTGTCTACGAGATGAGAT	600
QY	377	GLYTHRVALASNLEUVALSERLEULEUALALEUGINSERLEUTRIPLYS	392
Db	601	GGGACGTGTAACATGAGCGAGCTTGTGGCTTTGTGTGATGATGAGTGAGG	648

BASE COUNT 167 a 153 c 168 g 160 t

ORIGIN

Alignment Scores:

Pred. No.: 5.03e-90 Length: 648

Score: 809.00 Matches: 149

Percent Similarity: 83.80% Conservative: 32

Best Local Similarity: 68.98% Mismatches: 35

Query Match: 34.68% Indels: 0

De: 14 Gaps: 0

US-09-857-612A-14 (1-443) x BQ852109 (1-648)

/note="Vector: pBRCNDNA511AB. The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpb.ucdavis.edu/TAG_LIB-Q6_ABCDI lettuce salinas TAG_TISSUE-chemical TAG_SEQ-TGAGCCGGC"

RESULT 5 BQ855578 630 bp mRNA linear EST 14-AUG-2002

LOCUS BQ855578.1 y.g.ab1 Q6_ABCDI lettuce salinas lactuca sativa cDNA clone

DEFINITION

648

QGB27B05, mRNA sequence.
 BO855578
 VERSION BO855578.1 GI:22241043
 KEYWORDS
 SOURCE Lactuca sativa.
 ORGANISM Lactuca sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; eusterids II; Asterales; Asteraceae; Lactuceae; Lactuca

REFERENCE
 1 (bases 1 to 630)
 Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison,P., Koljman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.
 Lettuce and Sunflower ESTs from the Composite Genome Project
 http://compgenomics.ucdavis.edu/
 Unpublished (2002)
 Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@ucdavis.org [michelmoreveg@mail.ucdavis.edu]
 belongs to contig QG_CA.Contig973, see http://cgpbdb.ucdavis.edu/ for details.
 Plate: QGB27 row: B column: 05.

FEATURES
 source
 Location/Qualifiers
 1..630
 /organism="Lactuca sativa"
 /cultivar="Salinas"
 /db_xref="taxon:4236"
 /clone="QGB27B05"
 /clone_1fb="QG_ABCDI lettuce salinas"
 /lab_host="E.coli"
 /note="Vector: pBRCDNA5f1AB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpbdb.ucdavis.edu/
 TAG_LIB-QG_ABCDI lettuce salinas
 TAG_TISSUE-chemical induction
 TAG_SEQ-TGTAGCCGGC"

BASE COUNT 166 a 153 c 157 g 154 t
 ORIGIN

Alignment Scores:
 Pred. No.: 6,35e-87 Length: 630
 Score: 784.00 Matches: 145
 Percent Similarity: 83.81% Conservative: 31
 Best Local Similarity: 69.05% Mismatches: 34
 Query Match: 33,608 Indels: 0
 Db: 14 Gaps: 0

US-09-857-612a-14 (1-443) x BO855578 (1-630)

QY 177 GYHISProSerGIInValIGySerLyPheLeuLysAspLeuLysAsnLeuIleGlu 196
 Db 1 GGGCATCTCTGGCGACGCGGCTCCACCTCTCCAAACATTAACGATTAAATCGAGGAA 60

QY 197 ALASerAsnSerAsnAngLyLysProValIleLeuLeuSerHisSerLeuGlyGlyLeu 216
 Db 61 GCAAGGATTCATCAATGGTGGAAACCGCATCTCTATCTCCCATATCTTCGCGGCTC 120

QY 217 PheValLeuGluLeuLeuAsnArGAsnProProSerTrpArgLyLysPheIleLysHis 236
 Db 121 TTGCTCTCCACTCTCAACCGCAACCCACGCTCTGCGGAAACATATACATTAAGCAC 180

QY 237 PheIleAlaLeuSerAlaProTrpGlyGlyAlaIleAspGluMetCysTrpPheAlaSer 256
 Db 181 TTGATCGCATAGCGGACCATGGGGTGGACGGTTACGAGATGTGACTTTGCATGCC 240

QY 257 GLYAsnThrLeuGlyAlaProLeuValAspProLeuLeuValArgAspGluGlnArgSer 276
 Db 241 GGGAAATTCACAGAGAGGCCCTCCCTGTAACCCCTTGCTGTGTAGAAACGAACACGAGC 300

QY 277 SerGluSerLysLeuTrpLeuLeuProAsnProLysIlePheGlyProGlnLysProIle 296
 Db 301 TCAGAGACCACTTATGCTATTCGACATCAACAGAGATTCGCCGAGAGAAACCACTT 360

QY 297 ValIleThrProIleArgProTyrSerAlaHisAspMetValAspPheLeuLysAspIle 316
 Db 361 GTTGTTCACAAACACCTGCAGCTATCTCTTTGATATTCATATTCATATTTGAAGATATC 420

QY 317 GlyPheProGluGlyAlaTyrProTyrGluThrArgIleLeuProLeuIleGlyAsnIle 336
 Db 421 GCGTTCTCTGAAGGGGTGACGCTTACAGAAACGCGCTTGCCCTTGTTGAGAGAGTTG 480

QY 337 LysAlaProGluValProIleThrCysIleMetGlyThrGlyValGlyThrLeuGluThr 356
 Db 481 GCTGACACGAGAAATTCGATACATGATATTCGAAAGTGGAGTCAAGCCGGAATCT 540

QY 357 LeuPheTrpGlyLysGlyAspPheAspGluArgProGluLysSerTyrGlyAspGlyAsp 376
 Db 541 TTGCATTACGAGAGAAACGTTTGATATACGACCGGAGATTTGCTACGAGATGGAGAT 600

QY 377 GlyThrValAsnLeuValSerLeuLeuAla 386
 Db 601 GGCAGCGTGAAATGCGGAGCTGTGGCT 630

RESULT 6
 LOCUS BM85973
 DEFINITION semi0d11.y1 Gm-cl063 glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl063-5158 5' similar to TR:Q92W42 Q92W42 F17L21.28. ;, mRNA sequence.

ACCESSION BM85973
 VERSION BM85973.1 GI:19269717
 KEYWORDS soybean.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; euroids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE
 1 (bases 1 to 507)
 Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corvelli,V., Rhana,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritzer,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
 Public soybean EST project
 Unpublished (1999)
 Contact: Shoemaker R./Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com

TITLE
 JOURNAL
 COMMENT
 Seq primer: -40RP from Glibco
 High quality sequence stop: 421.
 Location/Qualifiers
 1..507
 /organism="Glycine max"
 /db_xref="taxon:3847"

FEATURES
 source

```

/clone="SOYBEAN CLONE ID: Gm-cl063-5158"
/clone.lib="Gm-cl063"
/tissue.type="Germinating shoot, 24 hour germination"
/lab.host="DH10B"
/note="vector: pluescript II SK+; Site_1: EcoRI, Site_2:
XhoI; The cDNA library was constructed from mRNA isolated
from germinating shoots of the cultivar Williams. The
seeds were allowed to germinate for 24 hours prior to
harvesting the germinating shoots. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (Gibco BRL). This
library was constructed in the laboratory of Dr. Randy
Shoemaker at Iowa state university."

```

Alignment Scores:	
Pred. No.:	7,83e-66
Score:	774.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	33.18%
DB:	14
	Gaps:
	0
US-09-657-612A-14 (1-443) x BM685973 (1-507)	

OY	1	MettysLysGluInGIuGIuGILeuLysIleGIuValAlaThrLeuThrValThrVal	20
Db	76	ATGAAGAAGCAACAAGAAGAGGTCTCAAGATTGACGTTCACACTCACATTACAGTA	135
OY	21	ValValValMetLeuSerLeuLeuCysThrCysGIuAlaSerAsnLeuaspProLeuile	40
Db	136	GTTTGTGGTAGTGCTTCTCATTCCTATGCACATGTGGGGCAAGCAACCTCGACCCTTTGATT	195
OY	41	LeuilePProGJLYasnGLIyAsnInleuGIuAlaArgLeuThrAsnGlnTrpLysPro	60
Db	196	CTAATACCAAGGTAAcGGAGGGCAcCACTAGAAAGAGGTGACCAATTCAGTACAAgCCC	255
OY	61	SerThrPheileCysGluSerTrpTyrrProLeuileIleLysLysAsnGIlyrrPhearg	80
Db	256	TCTACTTTCATCTGGAGATCATGTGTAccCTTCATCAcAAGAAAAGATGAGTGTTCAGA	315
OY	81	LeutrrPheaspSerSerValIIeLeuAlaProPhethrInCysPheAlaGIuarGmet	100
Db	316	CTTGGTGTTCATTCACAGTGCATACCTTCCTCTTCACACTCAATGCTTGGCCAAAGCAGTG	375
OY	101	ThrLeuHISyrHisGluInGIuleuAspAspTYrrPheasnthPrPGIValGIuThrarg	120
Db	376	ACCCTTCATTACACACAAAGAACTGCATGATTAATTCAACACACTCTGGGGTTAGAACCCGG	435
OY	121	ValProHISpheGJySerThrAsnSerLeuLeuTYrrLeuAsnProArgLeuLysHisIle	140
Db	436	GTCCTCACTTGGGTTCACCAACACTCTGTCTCATTCATCAATCTCGTCTCAAGCATATC	495
OY	141	ThrGIyTYrrMet 144	
Db	496	ACCGGATTCATG 507	

LOCUS B33221.181 673 bp mRNA linear EST 14-JUL-2000
DEFINITION NF010H12IN1F101 Insect herbivory Medicago truncatula cDNA clone
ACCESSION NF010H12IN 5', mRNA sequence.
VERSION B33221.1
KEYWORDS B33221.1 GI:9195958
SOURCE EST.
ORGANISM *Medicago truncatula*
Bartel medic.

REFERENCE	TITLE	JOURNAL	COMMENT
Enkaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta, Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae; eucotsids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.			
1 (bases 1 to 673)			
Korth, K., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Imman, J.T., Weller, J.W. and May, G.D.			
Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula Insect herbivory library			
Unpublished (2000)			
Contact: Korth K			
Dept. of Plant Pathology			
University of Arkansas			
217 Plant Science Building, Fayetteville, AR 72701, USA			
Tel: 501 575 5191			
Fax: 501 575 7601			
Email: k.korth@comp.uark.edu			
Medicago Genome Initiative accession: MGI:S:24007			
Insert length: 673	Std Error: 0.00		
Plate: 010	row: H	column: 12	
Seq primer: TCACACAGGAACAGCTATGAC.			

FEATURES	SOURCE
location/Qualifiers	1..673
/organism="Medicago truncatula"	
/db_xref="taxon:3880"	
/clone="NF010H12IN"	
/clone_lib="Insect herbivory"	
/tissue_type="local and systemic leaves"	
/dev_stage="mature"	
/note="Vector: Lambda Zap; Library was produced from fully expanded M. truncatula leaves of plants fed upon by Spodoptera exigua (beet armyworm) for 24 hours. Systemic undamaged leaves from injured plants) and wounded leaves were harvested and pooled."	
BASE COUNT	207 a 143 c 110 g 213 t
ORIGIN	

Alignment Scores:		
Pred. No.:	6,99e-82	Length: 673
Score:	744.00	Matches: 146
Percent Similarity:	84.24%	Conservative: 25
Best Local Similarity:	71.92%	Mismatches: 22
Query Match:	31.89%	Indels: 10
DB:	10	Gaps: 3
US-09-857-612A-14 (1-443) x BE322181 (1-673)		
QY	5 GlnGlnGlnGlyLeuLysIleGluValAlaThrValThrValIvalValMet	24
DB	70 AAGGACATAGCTCTCAAAATTTTGCATT--ACCAATACATTGAGACA--CTGTAGTATG	123
QY	25 LeuSerLeuLeuCys-----ThrCysGlyAlaSerAsnLeuAspPro	38
DB	124 GTATCGTAAATTTTGTTCACAAATGTGGCCCTGTAGTACCAACCAACAAACCTCAATCA	183
QY	39 LeuIleLeuIleProGlyAsnGlyGlyAsnGlnLeuGluAlaArgLeuThrAsnGlnTrp	58
DB	184 GTGATTAATAATCCAGGATGACGAGGAGACCACTAGACAGCAAACTATACCACTAAATAT	243
QY	59 LysProSerThrPheIleCysGlnSerTrpTrpProLeuIleLysLysAsnGlyTrp	78
DB	244 AAACCTTCTACCTTAAATTTGTGATCCATGTCACCTCTCTTCAAGAGAAAAATGGCTGG	303
QY	79 PheArgLeuTrpPheAspSerSerValIleLeuAlaProPheTrpGlnCysPheAlaGlu	98
DB	304 TTATGATTAATGTTGATTCACAAGTGTCTACTTCTCTCCCTTCAACCAAAATGTTTGGCTTCT	363
QY	99 ArgMetThrLeuIleIstYrHisGlnGluIleuAspAspArgPheAsnTrpProGlyValGlu	118
DB	364 CGCATGACTCTTTATTTGATGATCAAGACCTTGATATTACTTAATGTCTTCGAGGTGAG	423
QY	119 ThrArgValProHisPheGlySerThrAsnSerLeuLeuTrpLeuAsnProArgLeuLys	138

Db 424 ACACGTCCTCAGCTTTGGTCTACTTCTCTCTTATCTTAATCCTGCTTTAG 483
 Oy 139 Hs11eThrGlyTyrMetAlaProLeuValAspSerLeuGlnLysLeuGlyTyrAlaAsp 158
 Db 484 CTTGTACACAGGATCATGGCCCTTAGTATGATCATTAGAACAGCTTGGTTACATGAT 543
 Oy 159 GlyGluThrLeuPheGlyAlaProTyrAspPheArgTyrGlyLeuAlaAlaGluGlyHis 178
 Db 544 GGTCAAAATCTGTTGGAGCTCCCTATGATTTTCGATATGCTAGACAGCACCAGTCAC 603
 Oy 179 ProSerGlnVal-GlySerLysPheLeuLysAspLeuLysAsn-LeuIleGluGluAla 138
 Db 604 CCATCTCAAGTTGGTTCCTCAAAATTCCTAAACGACCTAAAGATTTCATGAAAAACCA 663
 Oy 198 erasn 199
 Db 664 GCAC 668
 RESULT 8
 BG523648
 LOCUS 587 bp mRNA linear EST J6-NOV-2001
 DEFINITION 34-12 Stevia field grown leaf CDNA Stevia rebaudiana CDNA 5', mRNA
 sequence.
 ACCESSION BG523648
 VERSION BG523648.1 GI:16947068
 KEYWORDS EST.
 SOURCE Stevia rebaudiana.
 ORGANISM Stevia rebaudiana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; easterids II; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Stevia.
 1 (bases 1 to 587)
 Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.
 Leaf ESTs from Stevia rebaudiana: A resource for gene discovery in
 diterpene synthesis
 Unpublished (2001)
 Contact: Jim Brandle
 Genomics and Biotechnology
 Agriculture and Agri-Food Canada - SCFRC
 1391 Sandford St., London, Ontario, CANADA, N5V 4T3
 Tel: 519 457 1470
 Fax: 519 457 3997
 Email: brandleje@agr.ca
 Seq primer: T3 promoter primer.
 Location/Qualifiers
 1..587
 /organism="Stevia rebaudiana"
 /strain="751/1501"
 /cultivar="Landrace"
 /db_xref="taxon:55670"
 /clone_lib="Stevia field grown leaf CDNA"
 /tissue_type="leaf"
 /dev_stage="field grown, mid-size"
 /lab_host="E. coli strain XL0LR"
 /note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI; This
 cDNA library was constructed from polyA+ enriched mRNA
 from field grown leaves. Mid-size actively growing leaves
 were collected and pooled from several plants and frozen
 immediately after harvesting in liquid nitrogen. The cDNA
 was prepared using an XhoI-poly(dT) linker-primer. An
 EcoRI adapter was ligated to the blunt end cDNA and the
 products were digested with EcoRI and XhoI enabling
 directional cloning into the lambda ZAP Express vector.
 The library was amplified using the host strain XL1-Blue
 MRF'. Mass excision of the library was performed to
 obtain pBK-CMV phagemid clones in the host strain XL0LR.
 Single pass DNA sequencing was performed using the T3
 promoter primer: 5' ATTAACCTCACTAAAGCA 3'. This library
 was constructed by Alex Richman."

Alignment Scores:
 Pred. No.: 9, 72e-76 Length: 587
 Score: 694.00 Matches: 132
 Percent Similarity: 82.29% Conservative: 26
 Best Local Similarity: 68.75% Mismatches: 33
 Query Match: 29.75% Indels: 1
 DB: 12 Gaps: 0

US-09-857-612A-14 (1-443) x BG523648 (1-587)

Oy 224 ArgAsnProSerThrPArgLysLysPheIleLysHisPheIleAlaLeuSerAlaPro 243
 Db 3 CGTAACCCACCCCTCATGGCCCAAAACATPACATTAATTAATTCGGCTTCCGCACCA 62
 Oy 244 ThrGlyGlyAlaIleAspGluMetTyrThrPheAlaSerGlyAsnThrLeuGlyValPro 263
 Db 63 TGGGTGTGAACGGTTGACGAGATGTTAACTTTGCTTCTGGGAACACTGTGGAGTTCCA 122
 Oy 264 LeuValAspProLeuLeuValArgAspGluGlnArgSerSerGluSerAsnLeuTrpLeu 283
 Db 123 CTGGTGAACCCGTTGCTGTCAGAAATGACCAACGAGTTCTGAAACATTTATGGCTC 182
 Oy 284 LeuProAsnProLysIlePheGlyProGlnLysProIleValIleThrProIleArgPro 303
 Db 183 ATGCCACGACGAAACAAATTCCTCAACAGACACTGTAGTTCACATCAGATTCAACT 242
 Oy 304 TyrSerAlaHisAspMetValAspPheLeuLysAspIleGlyPheProGluGlyValTyr 323
 Db 243 TACTCGCTCTTTGACATTCACGTTCTTAAAGATATCGGGTTCGAGAGGGGTGTCCAC 302
 Oy 324 ProTyrGluThrArgIleLeuProLeuIleGlyAsnIleLysAlaProGlnValProIle 343
 Db 303 CTTACGACGACGACAACTCTGCTGTCAGAAATGACCAACGAGTTCTGAAACATTTATGGCTC 362
 Oy 344 ThrCysIleMetGlyThrGlyValGlyThrLeuGluThrLeuPheTyrGlyLysGlyAsp 363
 Db 363 ACGGTATAGTGGAAACGGATGTCGCCGACACCGGAACCTTTGTATATGGGACGACGGG 422
 Oy 364 PheAspGluArgProGluIleSerTyrGlyAspGlyLysPheThrValAsnLeuValSer 383
 Db 423 TTGCATAGACGACCGGAATGTTATGAGATGCTACGAAACGGATATGCTAGT 482
 Oy 384 LeuLeuAla-LeuGlnSerLeuTrpLysGluLysAsnGluThrLeuValValVal 403
 Db 483 TTGTGGCTTTTGGACGACGACGATGAGATGAAAGATCAGCAATTTGAAGGGATTA 542
 Oy 403 sIleAspGlyValSerHisThrSerIleLeuLys 414
 Db 543 GCTTACNGGATTTCTCATACAGACATCTTAA 576
 RESULT 9
 AV441327/c 618 bp mRNA linear EST 14-NOV-2000
 LOCUS AV441327 Arabidopsis thaliana above-ground organ two to six-week
 DEFINITION old Arabidopsis thaliana CDNA clone AB242902.f 3', mRNA sequence.
 ACCESSION AV441327
 VERSION AV441327.1 GI:7611713
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 618)
 Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized and
 size-selected cDNA libraries
 DNA Res. 7, 175-180 (2000)
 JOURNAL MEDLINE 20363093
 COMMENT Contact: Erika Asamizu
 The first laboratory for Plant Gene Research
 Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizuekazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
Location/Qualifiers

1..618
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone_id="Arabidopsis thaliana above-ground organ two to six-week old"
/issue_type="aboveground organs"
/dev_stage="two to six-week old"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 163 a 155 c 125 g 175 t

ORIGIN

Alignment Scores:

Pred. No.: 3.3e-70 Length: 618
Score: 650.00 Matches: 127
Percent Similarity: 75.49% Conservative: 27
Best Local Similarity: 62.25% Mismatches: 42
Query Match: 27.86% Indels: 8
DB: 10 Gaps: 2

US-09-857-612a-14 (1-443) x AV41327 (1-618)

QY 228 seTTPaTgLySLySPheLIeLysHISpHeLIeAlaLeuSerAlaProTPrGlyGlyAla 247

DB 618 TCATGGCGCCCGCCAGTACATCAACACTTCTTGCACCTCGCCGCGCATGGGGTGGAGC 559

QY 248 ILAspGLeuTyrTrhPheAlaSerGlyAsnTrhLeuGlyValProLeuValAspPro 267

DB 558 ATCTCTCAGATGAGACATTTGCTTGGCAACACACTCGGTGTCCTTAGTTAACCTT 499

QY 268 LeuLeuValArgAspGluArgSerSerGluSerAsnLeuTrpLeuLeuProAsnPro 287

DB 498 TTGGTGGTCACAGCGCATCAGAGAGCTCCGAGAGTAACAATGGCTCTTCACATCAC 439

QY 288 LysILePheGlyProGln---LysProILeValILeTrhProILeArgProTyrSerAla 306

DB 438 AAAGCTTTTCACAGACAGAACACTAACCCGTTCTCGTAACCTCCCGAGTTAACTACACACT 379

QY 307 HIsAspMetValAspPheLeuLysAspILeGlyPheProGluGlyValTyrProTyrGlu 326

DB 378 TACGAGATGATCGGTTTGTTCAGACATTCGATTCTCAAGAGAGTTGGCTTACAAAG 319

QY 327 ThIaTgILeLeuProLeuILeGlyAsnILeLysAlaProGlnValProILeThCysIle 346

DB 318 ACAAGAGTGGCTTTTAAACAGAGAGCTGATGACTCCGGAGAGTCCAGTCACTTGCAAT 259

QY 347 MetGlyTrhGlyValGlyTrhLeuGlyLysPheTrhLeuPheTyrGlyLysGlyAspPheAspGlu 366

DB 258 TATGGGAGAGAGAGTTCATACACCGAGGTTTGCATGTTGAAAAAGAGATTCATACAG 199

QY 367 ArgProGluILeSerTyrGlyAspGlyAspGlyTrhValAsnLeuValSerLeuLeuAla 386

DB 198 CAACGAGAGATTAAGTATGAGATGAGATGAGAGCGGTTAATTTGGCAGACTTACACT 139

QY 387 LeuILeInsLeuTrpLysGluGlyLysAsnGlnTrhLeuLysValValLysIleAspGly 406

DB 138 TTG-----AAAGTCAATAGCTTGAACACCGTGAAGATTCATGAGA 100

QY 407 ValSerHISTrhSerIleLeuLysAspGlyValAlaLeuAsnGlnIleValGlyLysIle 426

DB 99 GTTTCGCGATCATCTATTAAGACGAGATTCGACATTAAGAGATTATGACAGACTT 40

QY 427 ThIaSerIleAsn 430

DB 39 TCAATTAATTAT 28

RESULT 10

BG522086

LOCUS BG522086 635 bp mRNA linear EST 01-FEB-2002

DEFINITION 18-18 Stevia field grown leaf cDNA Stevia rebaudiana cDNA 5', mRNA

ACCESSION BG522086

VERSION BG522086.1 GI:18465150

KEYWORDS EST.

SOURCE Stevia rebaudiana.

ORGANISM Stevia rebaudiana.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; easterids II; Asterales; Asteraceae; Asteroideae;

Helianthae; Stevia.

1 (bases 1 to 635).

Brandl, D.E., Richman, A., Swanson, A.K. and Chapman, B.P.

Leaf ESTs from Stevia rebaudiana: A resource for gene discovery in

diterpene synthesis

Unpublished (2001)

Contact: Jim Brandl

Genomics and Biotechnology

Agriculture and Agri-Food Canada - SCPPRC

1391 Sandford St., London, Ontario, CANADA, N5V 4T3

Tel: 519 457 1470

Fax: 519 457 3997

Email: brandlje@em.agr.ca

Seq primer: 73 promoter primer.

Location/Qualifiers

source

1..635

/organism="Stevia rebaudiana"

/strain="751/1501"

/cultivar="landrace"

/db_xref="taxon:55670"

/clone_id="Stevia field grown leaf cDNA"

/issue_type="leaf"

/dev_stage="field grown, mid-size"

/lab_host="E. coli strain XL0R"

/note="Vector: pBK-CMV. Site_1: EcoRI; Site_2: XhoI; This

cDNA library was constructed from polyA+ enriched mRNA

from field grown leaves. Mid-size actively growing leaves

were collected and pooled from several plants and frozen

immediately after harvesting in liquid nitrogen. The cDNA

was prepared using an XhoI-poly(dT) linker-primer. An

EcoRI adapter was ligated to the blunt end cDNA and the

products were digested with EcoRI and XhoI enabling

directional cloning into the lambda ZAP Express vector.

The library was amplified using the host strain XL1-Blue

MRP. Mass excision of the library was performed to

obtain pBK-CMV phagemid clones in the host strain XL0R.

Single pass DNA sequencing was performed using the 73

promoter primer: 5' ATTACCTCCTCAAGGCA 3'. This library

was constructed by Alex Richman.

BASE COUNT 158 a 157 c 164 g 155 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 6.08e-67 Length: 635
Score: 624.00 Matches: 121
Percent Similarity: 73.46% Conservative: 34
Best Local Similarity: 57.35% Mismatches: 44
Query Match: 26.75% Indels: 13
DB: 12 Gaps: 1

US-09-857-612a-14 (1-443) x BG522086 (1-635)

QY 11 ILleGValAlaTrhLeuTrhValTrhValValValMetLeuSerLeuLeuCysThr 30

DB 4 GTTCGGTGAAGCAGATACATGCGCTATCGCCAGATGATGCGGTGGCCACCATACACT 63

QY 31 CysGlyAlaSer-----AsnLeuAspPro 38

DB 64 TGTGCGGCAAGACATCATATACCATCACCACGACGACGAGAACCATCGGCTGATCCG 123

QY 39 LeuILeLeuILeProGlyAsnGlyLysAsnGlnLeuGluAlaArgLeuThAsnGlnTyr 58

Db 124 GTGATCTGTCGCTGGAGCGGTGGGAACACAGCTAGAACCCGGTTACGGCCAGAT 183
 QY 59 LysProSerThrphelecygsluserTPryProleuilelyslslysaanglytrp 78
 Db 184 AAGGGACATAGCTGCTGTGCAACCGCTTACCCGCTCAAGAAAGATGAGGGGGGTG 243
 QY 79 PheArgLeuThrPhePheSerSerValleLeuAlaProPheThrGlnCysPheAlaGlu 98
 Db 244 TTGAGGCTATGCTTCGAGGTCGGCTGTGTTCGACACATTGACCGAGTCTTTGGCGAT 303
 QY 99 ArgMetThrLeuHisThrHisGlnGluLeuAspAspTyrPheAsnThrProGlyValGlu 118
 Db 304 CGTATACCGCTTACTACTACCCAGATGTCGATGATTACAAATATGCTCCGAGATGAG 363
 QY 119 ThrArgValProHisPheGlySerThrAsnSerLeuLeuLeuLeuAsnProAlaGlyLeu 138
 Db 364 ACTAGAGTCTCAGTTGCTTCTACTCAATCCCTTCTACTGATCCTTCTTCAAG 423
 QY 139 HisLeThrGlyTyrMetAlaProLeuValAspSerLeuGlnLysLeuGlyTyrAlaAsp 158
 Db 424 CATATACAAACATACATGACGACATTGTCGATCTATACACACTGCTCAAAAGAT 483
 QY 159 GlyGluThrLeuPheGlyAlaProTyrAspPheArgTyrGlyLeuAlaGluGlyHis 178
 Db 484 ACCGAGAACTCTTTCGAGCTCATATGATTTTCGNTACGCGTTAGCATCCGAAGCAT 543
 QY 179 ProSerGlnValGlySerLysPheLeuLysAspLeuLysAsnLeuIleGluAlaSer 198
 Db 544 CCTTGATGTCGCTCACTTACCTCCAAACCTTAAGAGTTAATCGAAAGCGAAGC 603
 QY 199 AsnSerAsnAsnGlyLysProValIleLeuLeu 209
 Db 604 AC-ACAAATGGCGGACTCCCGGTGATCTCTGTC 635
 RESULT 11
 BG525847 596 bp mRNA linear EST 16-NOV-2001
 LOCUS 54-55 Stevia field grown leaf cDNA Stevia rebaudiana cDNA 5', mRNA
 DEFINITION
 sequence.
 ACCESSION BG525847
 VERSION BG525847.1 GI:16949308
 KEYWORDS EST.
 SOURCE Stevia rebaudiana.
 ORGANISM Stevia rebaudiana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; easterids II; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Stevia.
 1 (bases 1 to 596)
 Brande, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.
 Leaf ESTs from Stevia rebaudiana: A resource for gene discovery in
 diterpene synthesis
 Unpublished (2001)
 CONTACT: Jim Brande
 Genomics and Biotechnology
 Agriculture and Agri-Food Canada - SCPPRC
 1391 Sandford St., London, Ontario, CANADA, N5V 4T3
 Tel: 519 457 1470
 Fax: 519 457 3997
 Email: brandeje@em.agr.ca
 Seq primer: r3 promoter primer.
 Location/Qualifiers
 1..596
 /organism="Stevia rebaudiana"
 /strain="751/1501"
 /cultivar="Landrace"
 /db_xref="taxon:55670"
 /clone.lib="Stevia field grown leaf cDNA"
 /tissue_type="leaf"
 /dev_stage="field grown, mid-size"
 /lab_host="E. coli strain XL0R"
 /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI; This
 cDNA library was constructed from polyA+ enriched mRNA

from field grown leaves. Mid-size actively growing leaves
 were collected and pooled from several plants and frozen
 immediately after harvesting in liquid nitrogen. The cDNA
 was prepared using an XhoI-poly(dT) linker primer. An
 EcoRI adapter was ligated to the blunt end cDNA and the
 products were digested with EcoRI and XhoI enabling
 directional cloning into the lambda Zap Express vector.
 The library was amplified using the host strain XL1-Blue
 MR'. Mass excision of the library was performed to
 obtain pBK-CMV phagemid clones in the host strain XL0R.
 Single pass DNA sequencing was performed using the r3
 promoter primer: 5' ATTAACCTCACTAAAGGA 3'. This library
 was constructed by Alex Richman."

BASE COUNT 153 a 147 c 149 g 147 t
 ORIGIN

Alignment Scores:
 Pred. No.: 9,7e-66 Length: 596
 Score: 614.00 Matches: 114
 Percent Similarity: 73.74% Conservative: 32
 Best local Similarity: 57.58% Mismatches: 40
 Query Match: 26.32% Gaps: 12
 DB: 12 Gaps: 1

US-09-857-612a-14 (1-443) x BG525847 (1-596)

QY 15 ThrLeuThrValThrValValValMetLeuSerLeuLeuCysThrGlyAlaSer 34
 Db 3 ACGATACAGATGGCTATCGCAGCATGTGCGGTTGACCATATACCTTGGCGAAGC 62
 QY 35 -----AsnLeuAspProLeuIleLeuIle 42
 Db 63 ATCATATATACCATCACACACACACACACATGCGCTGATCCGGTATCTGTG 122
 QY 43 ProGlyAsnGlyGlnGlnLeuGluAlaArgLeuThrAsnGlnTyrLysProSerThr 62
 Db 123 CCTGAGACCGGTGGAGACCAAGCTGTAAGCCCGGTAAAGCGCGATTAAGGAGATACG 182
 QY 63 PheIleCysGlnSerTPryProLeuIleLysLysAsnGlyTyrPheArgLeuTrp 82
 Db 183 TGGTGTGCAACCGGTTCTACCCGCTCAAGAAAGATGAGGGGGTGTTCAGGCTATGG 242
 QY 83 PheAspSerSerValleLeuAlaProPheThrGlnCysPheAlaGluArgMetThrLeu 102
 Db 243 TTGAGAGTCGGCTGTGTGTGCAACATTGACGAGTCTTTGCCATTCGATACGCTT 302
 QY 103 HisTyrHisGlnGluLeuAspAspTyrPheAsnThrProGlyValGluThrArgValPro 122
 Db 303 TACTATGACCGAGATGTCGATGATTACAAATATGCTCCGAGAGGAGACTAGAGTGTCT 362
 QY 123 HisPheGlySerThrAsnSerLeuLeuTyrLeuAsnProArgLeuLysHisIleThrGly 142
 Db 363 CAGTTTGTCTTACTCAATGCTTCTTACCTTCTTCAAGCATATTAACACA 422
 QY 143 TyrMetAlaProLeuValAspSerLeuGlnLysLeuGlyTyrAlaAspGlyThrLeu 162
 Db 423 TACATGACACATTTGGTGGAGTCTATATACACACTTGGCTACAAAGATACCGAATCTT 482
 QY 163 PheGlyAlaProTyrAspPheArgTyrGlyLeuAlaGluGlnHisProSerGlnVal 182
 Db 483 TTGGAGCTCCATATGATTTTCGTACGGGTACATCCGAAGACATTCCTTGATGATGTC 542
 QY 183 GlySerLysPheLeuLysAspLeuLysAsnLeuIleGluAlaSerAsnSer 200
 Db 543 GCGTCAACTTACCTCCAAACCTTAAGCAGTTAATCGAAAGCGAAGCAACACT 596
 RESULT 12
 BG127829 724 bp mRNA linear EST 31-JAN-2001
 LOCUS EST473475 tomato shoot/meristem Lycopersicon esculentum cDNA clone
 DEFINITION
 cDNA library was constructed from polyA+ enriched mRNA
 ACCESSION BG127829
 VERSION BG127829.1 GI:12628017

KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;
Asteridae: euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 724)
AUTHORS van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Utterback, T.,
Hansen, C., Ronning, C. and Tanksley, S.
TITLE Generation of ESTs from tomato shoot/meristem tissue
JOURNAL Unpublished (2001)
COMMENT Contact: CUCI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.
FEATURES
source
1..724
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cT0F18A6"
/clone_lib="tomato shoot/meristem"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."
BASE COUNT 238 a 138 c 124 g 224 t
ORIGIN
Alignment Scores:
Pred. No.: 1,15e-62 length: 724
Score: 590.50 Matches: 126
Percent Similarity: 59.07% Conservative: 27
Best Local Similarity: 48.65% Mismatches: 45
Query Match: 25,31% Indels: 63
DB: 12 Gaps: 1
US-09-857-612a-14 (1-443) x BG127829 (1-724)
QY 11 TlGluValAlArhrLeuthrValThrValValValMetLeuSerLeuLeuCysThr 30
::: ::: ||| ::: : : : : : ||||| |||
Db 131 TTAGGCATGAAGTCCTAATCATATCATATATCAACAATGTTGTCACATATTGCCAA 190
QY 31 CysGlyAlaSerAsnLeuAspProLeuIleuIleProGlyAsnGlyAsnGlnLeu 50
::: ||||| : : : : : ||||| |||||
Db 191 GCAATTACTACTATCTCAACCCCTATCTTGATACAGAGCTGGTGAACCAATTA 250
QY 51 GluAlaArgLeuthrAsnGlnIleValProSerThrPheIleCysGlnSerTrpTrpPro 70
||||| ||||| : : : : : ||||| |||||
Db 251 GAGCAAGATTAACTTCAAGATACAGACCTTATGTCACAACTTTATATCCA 310
QY 71 LeuIleLeuLysLysAsnGlyTrpPheArgLeuTrpPheAspSerSerValIleLeuAla 90
||| ||| |||
Db 311 TTAAATGAAGATAGTGAAGAAA----- 333
QY 91 ProPheThrGlnCysPheAlaGluArgMetThrLeuHisTrpHisGlnGluLeuAspAsp 110
333 ----- 333
QY 111 TyrPheAsnThrProGlyValGluThrArgValProHisPheGlySerThrAsnSerLeu 130
333 ----- 333
Db 131 LeuTyrLeuAsnProArgLeuLysHisIleThrGlyTyrMetAlaProLeuValAspSer 150
||||| ||||| ||||| : : :
Db 334 -----ATTACAGCATATATGAGGACCACTTGTGAAGCT 366
QY 151 LeuGlnLysLeuGlyTyrAlaAspGlyGluThrLeuPheGluAlaProTyrAspPheArg 170

Db 367 TTGGAGAAATAGGCTATAGAGATGTCAAATCTGTTGGAGACCCCTACGATTTTCT 426
QY 171 TyrGlyLeuAlaAlaGluGlyHisProSerGlnValGlySerLysPheLeuLysAspLeu 190
||||| ||||| ||||| ||||| : : : : : ||||| |||||
Db 427 TATGTTGGCTGCGCAGAGTCATCAAGTCATGTTGGTTGCAATATCTCAGAGACTTG 486
QY 191 LysAsnLeuIleGluAlaSerAsnSerAsnGlyLysProValIleLeuLeuSer 210
||| ||||| ||||| ||||| : : : : : ||||| |||||
Db 487 AAAGCACTGATGAAGAAGCGAAGTAAATTCATTAAGAGGAGTAAACCGTAAATCTTGTCT 546
QY 211 HisSerLeuGlyGlyLeuPheValLeuIleLeuAsnArgAsnProProSerTrpArg 230
||||| ||||| ||||| ||||| : : : : : ||||| |||||
Db 547 AC-AGCTTAGGTGCTATATGTCGACACTACTGTTGCTAATCAAAATGCTGGAGT 605
QY 231 LysLysPheIleLysHisPheIleAlaLeuSerAlaProTrpGlyAlaIleAspGlu 250
::: ||||| ||||| ||||| ||||| : : : : : ||||| |||||
Db 606 CAAAGTACATCAACACATTTATACATTTGGCTTCACCATGGGAGGAGCAAGTTATGCA 665
QY 251 MetTyrThrPheAlaSerGlyAsnThrLeuGlyValProLeuValAspProLeuLeu 269
||| ||||| ||||| ||||| ||||| ||||| : : : : : ||||| |||||
Db 666 ATGCTACTTTTGCTTGTAACACACTTGTGCTTAAGTAATCCGTTACTC 722
RESULT 13
BG526525 616 bp mRNA Linear EST 16-NOV-2001
BG526525
DEFINITION 60-13 Stevia field grown leaf cDNA Stevia rebaudiana cDNA 5', mRNA
sequence.
ACCESSION BG526525
VERSION BG526525.1 GI:16950014
KEYWORDS EST.
SOURCE Stevia rebaudiana.
ORGANISM Stevia rebaudiana
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;
Asteridae: euasterids II; Asterales; Asteraceae; Asteroideae;
Helianteae; Stevia.
1 (bases 1 to 616)
AUTHORS Brande, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.
TITLE Leaf ESTs from Stevia rebaudiana: A resource for gene discovery in
diterpene synthesis
JOURNAL Unpublished (2001)
COMMENT Contact: Jim Brande
Genomics and Biotechnology
Agriculture and Agri-Food Canada - SCPPRC
1391 Sandford St., London, Ontario, CANADA, N5V 4T3
Tel: 519 457 1470
Fax: 519 457 3997
Email: brandeje@em.agr.ca
Seq primer: T3 promoter primer.
FEATURES
source
1..616
/organism="Stevia rebaudiana"
/strain="751/1501"
/cultivar="Landrace"
/db_xref="taxon:55670"
/clone_lib="Stevia field grown leaf cDNA"
/tissue_type="leaf"
/dev_stage="field grown, mid-size"
/lab_host="E. coli strain XL0R"
/note="Vector: pBK-CMV; Site.1: EcoRI; Site.2: XhoI; This
cDNA library was constructed from polyA+ enriched mRNA
from field grown leaves. Mid-size actively growing leaves
were collected and pooled from several plants and frozen
immediately after harvesting in liquid nitrogen. The cDNA
was prepared using an XhoI-poly(dT) linker-primer. An
EcoRI adapter was ligated to the blunt end cDNA and the
products were digested with EcoRI and XhoI enabling
directional cloning into the lambda ZAP Express vector.
The library was amplified using the host strain XL1-Blue
MR'. Mass excision of the library was performed to
obtain pBK-CMV phagemid clones in the host strain XL0R.
Single pass DNA sequencing was performed using the T3

Promoter primer: 5' ATTAACCTCTACTAAGGGA 3'. This library was constructed by Alex Richman.

BASE COUNT	168 a	117 c	162 g	161 t	8 others
------------	-------	-------	-------	-------	----------

Alignment Scores:

Pred. No.:	1.35e-60	Length:	616
Score:	573.00	Matches:	112
Percent Similarity:	80.24%	Conservative:	22
Best Local Similarity:	67.07%	Mismatches:	31
Query Match:	24.56%	Indels:	2
DB:	12	Gaps:	0

US-09-857-612A-14 (1-443) x BG526525 (1-616)

QY 224 ATGAAGPProSerThrPArgLysLysPheLleLysHisPheLleAlaLeuSerAlaPro 243
|||||
DB 3 CGTAAACCCACCTCATGGGCGCAACATTAACATTATTCGCGTNCGCCGACCA 62
QY 244 TTPGLYLAlAlLeaSPGLuMetTyrThrPheAlaSerGlyAsnThrLeuGly-ValPr 263
|||||
DB 63 TGGGTGGAGACGGTTGACAGATGTTACTTTGCTTGGGAAACACTCTGGGAAGTTCC 122
QY 263 OLeuValaSPProLeuLeuValaArgSPGLuGlnArgSerSerGluSerAsnLeuTrpLe 283
|||||
DB 123 ACTGTGAACCCGTTGCTGCTGCAAGATGACGAACGAGTTTGAAAGCCATTATGGCC 182
QY 283 uLeuProAsnProLysLlePheGlyProGlnLysProLleValLleThrProLleArgPr 303
|||||
DB 183 CATGCCACGACGAAACAAATTCCTCAACAGACACTGTGTCACATTCAGATTCAAC 242
QY 303 CTYrSerAlaHisASPmetValaSPheLeuLysAspLleGlyPheProGluGlyValTy 323
|||||
DB 243 TTAAGTGTCTTTTGACATTTCACGGTTTCTTAAGATATCGGTTGGAAGAGGGTGCA 302
QY 323 rProTyrGlnThrArgLleLeuProLeuLleGlyAsnLleLysAlaProGlnValProL 343
|||||
DB 303 CCTTACGAGACACGAAATCTTGCTTGGTGAAGATTGGTTGTCACCGGAGTGCCTG 362
QY 343 eThrCysLleMetGlyThrGlyAlaGlyThrLeuGlnThrLeuPheThrGlyLysGly 363
|||||
DB 363 AACGTCTAAGTGGAAACGGTGTCCGACACCGCAACCTTGTATTATGGGACGACGCG 422
QY 363 PheAspGluArgProGluLleSerTyrGlyAspGlyLysGlyValaSPLeuVal-S 383
|||||
DB 423 GTTCGTAAGCAGCCGAAATGTTTATGAGATGATGACGAGCGTGATATGGTAG 482
QY 383 eXLeuLeuAlaLeuGln 388
|||||
DB 483 GTTGTGGCTTTGGAG 499

RESULT 14
BE355801 537 bp mRNA linear EST 20-JUL-2000
LOCUS BE355801
DEFINITION DGI_11.E05.b1_A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA
sequence.

ACCESSION BE355801
VERSION BE355801.1 GI:9296903
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
Clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 537)
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudhan,M. and Pratt
,L.H.

TITLE An EST database from Sorghum: dark-grown seedlings
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860
Fax: 706 583 0210
Email: mmp@pratt.uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 460
POLYA=NO.

FEATURES

source

1..537
Location/Qualifiers
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DGI)"
/note="Organ: 5-day-old dark-grown seedlings; Vector:
lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
made from poly-A RNA in the cloning vector lambda Zap II.
Clones to be sequenced were prepared by mass excision."

BASE COUNT 63 a 201 c 194 g 79 t
ORIGIN

Alignment Scores:

Pred. No.:	3.37e-59	Length:	537
Score:	561.00	Matches:	104
Percent Similarity:	74.01%	Conservative:	27
Best Local Similarity:	58.76%	Mismatches:	44
Query Match:	24.05%	Indels:	2
DB:	10	Gaps:	2

US-09-857-612A-14 (1-443) x BE355801 (1-537)

QY 60 ProSerThrPheLleCysGluSerTrpTyrProLeuLleLysLysAsnGlyTrpPhe 79
|||||
DB 2 CCGTCAGCCTCGTGCGCGGTTGG--CCACTGTCGGCGGCCCGCGGCTGCTTC 58
QY 80 ArgLeuTrpPheAspSerSerValLleLeuAlaProPheThrGlnCysPheAlaGluArg 99
|||||
DB 59 CGCCTCGGTTCGACCCCTCGTGCTGCGCGCGGCTCACAGGTGCTTCGCGAGCGG 118
QY 100 MetThrLeuHisTyrHisGlnGluLeuAspAspTyrPheAsnThrProGlyValGluThr 119
|||||
DB 119 ATGACGCTCCACACGCGCCGCGGACGACTACCGCAACGCCGCGCGCTGCAGAAC 178
QY 120 ArgValProHisPheGlySerThrAsnSerLeuLeuTyrLeuAsnProArgLeuLysHis 139
|||||
DB 179 AGGGTCCGACCTTCGGCTCACCTCCAGCTTCGCTACCTGACCCCAACCTCAAGCTC 238
QY 140 IleThrGlyTyrMetAlaProLeuValaSPSerLeu--GlnLysLeuGlyTyrAlaAsp 158
|||||
DB 239 CTGACGGGTACATGAACACTCTGCGGACGACGCTGAGAGAGAAAGCGGCTACGAGGAG 298
QY 159 GlyGluThrLeuPheGlyAlaProTyrAspPheArgTyrGlyLeuAlaGluGlyHis 178
|||||
DB 299 GGGCGGACCTGTCGGCGCGCGCTACGACTTCGCTGACGCGGCGCGGCGGCGGCGAC 358
QY 179 ProSerGlnValGlySerLysPheLeuLysAspLeuLysAsnLeuLleGluGluAlaSer 198
|||||
DB 359 CCGTCGAGGTGGACAGCGCTGACAGCGGCTGAGGCTGAGGAGGTGGCGGCTGC 418
QY 199 AsnSerAsnGlyLysProValLleLeuLeuSerHisSerLeuGlyGlyLeuPheVal 218
|||||
DB 419 GCGGCGAAGCGGCGGCGCGCGGCGGAGTCTGTGGCGACAGCGGCGGCGGCTGTTCGCG 478
QY 219 LeuGlnLeuLeuAsnArgAsnProProSerTrpArgLysLysPheLleLys 235
|||||
DB 479 CTCAGCTGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 529
RESULT 15
B0864610 527 bp mRNA linear EST 14-AUG-2002
LOCUS B0864610
DEFINITION QGC27C22.yg.ab1 QC_ABCDI lettuce salinas lactuca sativa cDNA clone
QGC27C22, mRNA sequence.
ACCESSION B0864610

VERSION B0864610.1 GI:22250075
KEYWORDS EST.
SOURCE Lactuca sativa.
ORGANISM Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; easterids II; Asterales; Asteraceae; Lactuceae;
Lactuca.
REFERENCE 1 (bases 1 to 527)
AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
Lin, H., Van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison
P., Kolman, J., Slabugh, M.S., Livingston, K., Zhou, T., Lai, Z.,
Church, S., Jackson, L. and Bradford, K.
TITLE Lettuce and Sunflower ESTs from the Composite Genome Project
JOURNAL http://compgenome.ucdavis.edu/
COMMENT Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
singletton, see http://cgpb.ucdavis.edu/ for details.
Plate: QGC27 row: C column: 22.
Location/Qualifiers
1. 527
/organism="Lactuca sativa"
/cultivar="Salinas"
/db_xref="taxon:4236"
/clone="QGC27C22"
/clone_id="QG-ABCDI lettuce salinas"
/lab_host="E.coli"
/note="Vector: pBRCDNA5flab: The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpb.ucdavis.edu/
TAG_LIB-QG-ABCDI lettuce salinas
TAG_ISSUE-flowers post-fertilized
TAG_SEQ-TGCCATCGCG"

BASE COUNT 122 a 125 c 147 g 133 t
ORIGIN

Alignment Scores:
Pred. No.: 3.78e-59 Length: 527
Score: 560.50 Matches: 102
Percent Similarity: 76.47% Conservative: 28
Best local Similarity: 60.00% Mismatches: 33
Query Match: 24.02% Indels: 7
DB: 14 Gaps: 1

US-09-857-612a-14 (1-443) x B0864610 (1-527)

OY 11 IlleGlVAlAlAThrLeuThrValValValValMetLeuSerLeuLeuGlyThr 30
DB 16 GTCAATGTGGCCACCGCATGATGATTACTATGATGGTGGCTGGCCCTCCACC 75

OY 31 CysGlyAla-----SerAsnLeuAspProLeuIleLeuIlePro 43
DB 76 TGGCGGGGGAATATCAACGATACCAACCGGAGCTCTATCCGGTATGTTCTGTCCT 135

OY 44 GlyAsnGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 63
DB 136 GGAGCCCGCGGAGCAACTGAGAGCCGATTGACGCGGAGTAAAGGGAGCAGAGCTGG 195

OY 64 IleCysGlnSerTrpTyrProLeuIleLeuIleLeuIleLeuIleLeuIleLeuIle 83
DB 196 CTATGACGCCGATTCTACCGATTAGAGACGAGGAGGAGGAGGAGGAGGAGGAGG 255

OY 84 AspSerSerValIleLeuAlaProPheThrGlnCysPheAlaGluArgMetThrLeuHis 103
DB 256 GACGTGGGTGATATATAGACACATCTTGGAGTGTGGTGGAGCATATGACGCTTAC 315

OY 104 TyrHisGlnGlnLeuAspAspTyrPheAsnThrProGlyValGluThrArgValProHis 123
DB 316 TACCAACCGAGAAATGATGATTAACAGAAATGCTCCGAGATTGACAGAGGTGTGTCAG 375

OY 124 PheGlySerThrAsnSerLeuLeuTyrLeuAsnProArgLeuLeuHisIleThrGlyTyr 143
DB 376 TTTGGTTCCACTGACATCCCTTCTCTACCTCGATCTTAATCTACACATATACGAGTTAC 435

OY 144 MetAlaProLeuValAspSerLeuGlnLeuGlyTyrAlaAspGlyGluThrLeuPhe 163
DB 436 ATGAACCATTTGGTGGAGTCTCTACAGAGCTTGGCTACACGATACAAACCATGTTC 495

OY 164 GlyAlaProTyrAspPheArgTyrGlyLeu 173
DB 496 GGAGCACCCTTACGATTTCCGTTATGGCCTT 525

Search completed: February 15, 2003, 15:21:20
Job time : 1685 secs